

# Particle cosmology

A. Riotto

CERN, Geneva, Switzerland

## Abstract

In these lectures the present status of the so-called standard cosmological model, based on the hot Big Bang theory and the inflationary paradigm is reviewed. Special emphasis is given to the origin of the cosmological perturbations we see today under the form of the cosmic microwave background anisotropies and the large scale structure and to the dark matter and dark energy puzzles.

## 1 Introduction

The evolution of the universe is determined to a large extent by the same microphysics laws of physics that govern high-energy physics phenomena. Hence, any progress in particle physics has a large impact on the cosmological model(s) and, conversely, any new step taken towards the understanding of the past, present and future of our universe might provide a hint of high-energy physics beyond the one we currently know. This is the reason why these lectures are entitled Particle Cosmology. If the reader takes only one lesson home from them it is that particle physics and cosmology are nowadays intimately connected.

There are fundamental questions we are on the edge of answering: what is the origin of our universe? Why is the universe so homogeneous and isotropic on large scales? What are the origins of dark matter and dark energy? What is the fate of our universe? While these lectures will certainly not be able to give definite answers to them, we shall try to provide the students with some tools they might find useful in order to solve these overwhelming mysteries themselves.

These lectures will contain a short review of the standard Big Bang model; a rather long discussion of the inflation paradigm with particular emphasis on the possibility that the cosmological seeds originated from a period of primordial acceleration; the physics of the Cosmic Microwave Background (CMB) anisotropies, and a discussion of the dark matter and dark energy puzzles.

Since these lectures were delivered at a school, we shall not provide an exhaustive list of references to original material, but refer to several basic cosmology books and reviews where students can find the references to the original material [1–8].

## 2 Basics of the Big Bang model

We know two basic facts about our local universe (the universe we may observe). First, it is homogeneous and isotropic on sufficiently large cosmological scales [2]. Once this experimental evidence is accepted, one can promote it to a principle, dubbed “the cosmological principle”. Secondly, it expands. The next question would then be: how can we describe such a universe?

The standard cosmology is based upon the maximally spatially symmetric Friedmann–Robertson–Walker (FRW) line element

$$ds^2 = -dt^2 + a(t)^2 \left[ \frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2 \theta d\phi^2) \right] ; \quad (1)$$

where  $a(t)$  is the cosmic-scale factor,  $R_{\text{curv}} \equiv a(t)|k|^{-1/2}$  is the curvature radius, and  $k = -1, 0, 1$  is the curvature signature. All three models are without boundary: the positively curved model is finite and curves back on itself; the negatively curved and flat models are infinite in extent. The Robertson–Walker metric embodies the observed isotropy and homogeneity of the universe. It is interesting to note that

this form of the line element was originally introduced for the sake of mathematical simplicity; we now know that it is well justified at early times or today on large scales ( $\gg 10$  Mpc), at least within our visible patch.

The coordinates,  $r$ ,  $\theta$ , and  $\phi$ , are referred to as *co-moving* coordinates: A particle at rest in these coordinates remains at rest, i.e., constant  $r$ ,  $\theta$ , and  $\phi$ . A freely moving particle eventually comes to rest in these coordinates, as its momentum is redshifted by the expansion,  $p \propto a^{-1}$ . Motion with respect to the co-moving coordinates (or cosmic rest frame) is referred to as peculiar velocity; unless supported by the inhomogeneous distribution of matter, peculiar velocities decay away as  $a^{-1}$ . Thus the measurement of peculiar velocities, which is not easy as it requires independent measures of both the distance and velocity of an object, can be used to probe the distribution of mass in the universe.

Physical separations between freely moving particles scale as  $a(t)$ ; or said another way the physical separation between two points is simply  $a(t)$  times the coordinate separation. The momenta of freely propagating particles decrease, or redshift, as  $a(t)^{-1}$ , and thus the wavelength of a photon stretches as  $a(t)$ , which is the origin of the cosmological redshift. The redshift suffered by a photon emitted from a distant galaxy  $1 + z = a_0/a(t)$ ; that is, a galaxy whose light is redshifted by  $1 + z$ , emitted that light when the universe was a factor of  $(1 + z)^{-1}$  smaller. When the light from the most distant quasar yet seen ( $z = 4.9$ ) was emitted, the universe was a factor of almost six smaller; when CMB photons last scattered, the universe was about 1100 times smaller.

## 2.1 Friedmann equations

The evolution of the scale factor  $a(t)$  is governed by Einstein equations

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \equiv G_{\mu\nu} = 8\pi G T_{\mu\nu} \quad (2)$$

where  $R_{\mu\nu}$  ( $\mu, \nu = 0, \dots, 3$ ) is the Riemann tensor and  $R$  is the Ricci scalar constructed via the metric (1) [2], and  $T_{\mu\nu}$  is the energy-momentum tensor.  $G = m_{\text{Pl}}^{-2}$  is the Newton constant. Under the hypothesis of homogeneity and isotropy, we can always write the energy-momentum tensor under the form  $T_{\mu\nu} = \text{diag}(\rho, P, P, P)$  where  $\rho$  is the energy density of the system and  $P$  its pressure. They are functions of time.

The evolution of the cosmic-scale factor is governed by the Friedmann equation

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G \rho}{3} - \frac{k}{a^2}, \quad (3)$$

where  $\rho$  is the total energy density of the universe, matter, radiation, vacuum energy, and so on.

Differentiating wrt to time both members of Eq. (3) and using the the mass conservation equation

$$\dot{\rho} + 3H(\rho + P) = 0, \quad (4)$$

we find the equation for the acceleration of the scale factor

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P). \quad (5)$$

Combining Eqs. (3) and (5) we find

$$\dot{H} = -4\pi G(\rho + P). \quad (6)$$

The evolution of the energy density of the universe is governed by

$$d(\rho a^3) = -P d(a^3); \quad (7)$$

which is the first law of thermodynamics for a fluid in the expanding universe. (In the case that the stress energy of the universe is comprised of several, non-interacting components, this relation applies to each

separately; *e.g.*, to the matter and radiation separately today.) For  $P = \rho/3$ , ultra-relativistic matter,  $\rho \propto a^{-4}$  and  $a \sim t^{\frac{1}{2}}$ ; for  $P = 0$ , very nonrelativistic matter,  $\rho \propto a^{-3}$  and  $a \sim t^{\frac{2}{3}}$ ; and for  $P = -\rho$ , vacuum energy,  $\rho = \text{const.}$  If the rhs of the Friedmann equation is dominated by a fluid with equation of state  $P = w\rho$ , it follows that  $\rho \propto a^{-3(1+w)}$  and  $a \propto t^{2/3(1+w)}$ .

We can use the Friedmann equation to relate the curvature of the universe to the energy density and expansion rate:

$$\Omega - 1 = \frac{k}{a^2 H^2}; \quad \Omega = \frac{\rho}{\rho_{\text{crit}}}; \quad (8)$$

and the critical density today  $\rho_{\text{crit}} = 3H^2/8\pi G = 1.88h^2 \text{ g cm}^{-3} \simeq 1.05 \times 10^4 \text{ eV cm}^{-3}$ . There is a one-to-one correspondence between  $\Omega$  and the spatial curvature of the universe: positively curved,  $\Omega_0 > 1$ ; negatively curved,  $\Omega_0 < 1$ ; and flat ( $\Omega_0 = 1$ ). Further, the fate of the universe is determined by the curvature: model universes with  $k \leq 0$  expand forever, while those with  $k > 0$  necessarily recollapse. The curvature radius of the universe is related to the Hubble radius and  $\Omega$  by

$$R_{\text{curv}} = \frac{H^{-1}}{|\Omega - 1|^{1/2}}. \quad (9)$$

In physical terms, the curvature radius sets the scale for the size of spatial separations where the effects of curved space become pronounced. And in the case of the positively curved model it is just the radius of the 3-sphere.

The energy content of the universe consists of matter and radiation (today, photons and neutrinos). Since the photon temperature is accurately known,  $T_0 = 2.73 \pm 0.01 \text{ K}$ , the fraction of critical density contributed by radiation is also accurately known:  $\Omega_R h^2 = 4.2 \times 10^{-5}$ , where  $h = 0.72 \pm 0.07$  is the present Hubble rate in units of  $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$  [9]. The remaining content of the universe is another matter. Rapid progress has been made recently toward the measurement of cosmological parameters [10]. Over the past years the basic features of our universe have been determined. The universe is spatially flat; accelerating; comprised of one third dark matter and two thirds a new form of dark energy. The measurements of the cosmic microwave background anisotropies at different angular scales performed by the WMAP Collaboration [9] have recently significantly increased the case for accelerated expansion in the early universe (the inflationary paradigm) and at the current epoch (dark energy dominance), especially when combined with data on high-redshift supernovae (SN1) and large-scale structure (LSS) [10]. The CMB+LSS+SN1 data give [9]

$$\Omega_0 = 1.00^{+0.07}_{-0.03},$$

meaning that the present universe is spatially flat (or at least very close to being flat). Restricting to  $\Omega_0 = 1$ , the dark matter density is given by [9]

$$\Omega_{\text{DM}} h^2 = 0.11^{+0.0034}_{-0.0059},$$

and a baryon density

$$\Omega_B = 0.045 \pm 0.0015,$$

while the Big Bang nucleosynthesis estimate is  $\Omega_B h^2 = 0.019 \pm 0.002$ . Substantial dark (unclustered) energy is inferred:

$$\Omega_{\text{DE}} \approx 0.72 \pm 0.015.$$

What is most relevant for us is that this universe was apparently born from a burst of rapid expansion, inflation, during which quantum noise was stretched to astrophysical size seeding cosmic structure. This is exactly the phenomenon we want to address in part of these lectures.

## 2.2 The early, radiation-dominated universe

In any case, at present, matter outweighs radiation by a wide margin. However, since the energy density in matter decreases as  $a^{-3}$ , and that in radiation as  $a^{-4}$  (the extra factor due to the redshifting of the energy of relativistic particles), at early times the universe was radiation dominated—indeed the calculations of primordial nucleosynthesis provide excellent evidence for this. Denoting the epoch of matter and radiation equality by subscript ‘EQ,’ and using  $T_0 = 2.73$  K, it follows that

$$a_{\text{EQ}} = 4.18 \times 10^{-5} (\Omega_0 h^2)^{-1}; \quad T_{\text{EQ}} = 5.62 (\Omega_0 h^2) \text{ eV}; \quad (10)$$

$$t_{\text{EQ}} = 4.17 \times 10^{10} (\Omega_0 h^2)^{-2} \text{ s}. \quad (11)$$

At early times the expansion rate and age of the universe were determined by the temperature of the universe and the number of relativistic degrees of freedom:

$$\rho_{\text{rad}} = g_*(T) \frac{\pi^2 T^4}{30}; \quad H \simeq 1.67 g_*^{1/2} T^2 / m_{\text{Pl}}; \quad (12)$$

$$\Rightarrow a \propto t^{1/2}; \quad t \simeq 2.42 \times 10^{-6} g_*^{-1/2} (T / \text{GeV})^{-2} \text{ s}; \quad (13)$$

where  $g_*(T)$  counts the number of ultra-relativistic degrees of freedom ( $\approx$  the sum of the internal degrees of freedom of particle species much less massive than the temperature) and  $m_{\text{Pl}} \equiv G^{-1/2} = 1.22 \times 10^{19}$  GeV is the Planck mass. For example, at the epoch of nucleosynthesis,  $g_* = 10.75$  assuming three, light ( $\ll$  MeV) neutrino species; taking into account all the species in the Standard Model,  $g_* = 106.75$  at temperatures much greater than 300 GeV.

A quantity of importance related to  $g_*$  is the entropy density in relativistic particles,

$$s = \frac{\rho + p}{T} = \frac{2\pi^2}{45} g_* T^3,$$

and the entropy per co-moving volume,

$$S \propto a^3 s \propto g_* a^3 T^3.$$

By a wide margin most of the entropy in the universe exists in the radiation bath. The entropy density is proportional to the number density of relativistic particles. At present, the relativistic particle species are the photons and neutrinos, and the entropy density is a factor of 7.04 times the photon-number density:  $n_\gamma = 413 \text{ cm}^{-3}$  and  $s = 2905 \text{ cm}^{-3}$ .

In thermal equilibrium—which provides a good description of most of the history of the universe—the entropy per co-moving volume  $S$  remains constant. This fact is very useful. First, it implies that the temperature and scale factor are related by

$$T \propto g_*^{-1/3} a^{-1}, \quad (14)$$

which for  $g_* = \text{const}$  leads to the familiar  $T \propto a^{-1}$ .

Second, it provides a way of quantifying the net baryon number (or any other particle number) per co-moving volume:

$$N_B \equiv R^3 n_B = \frac{n_B}{s} \simeq (4 - 7) \times 10^{-11}. \quad (15)$$

The baryon number of the universe tells us two things: (1) the entropy per particle in the universe is extremely high, about  $10^{10}$  or so compared to about  $10^{-2}$  in the Sun and a few in the core of a newly formed neutron star. (2) The asymmetry between matter and antimatter is very small, about  $10^{-10}$ , since at early times quarks and antiquarks were roughly as abundant as photons. One of the great successes of particle cosmology is baryogenesis, the idea that  $B$ ,  $C$ , and  $CP$  violating interactions occurring out-of-equilibrium early on allow the universe to develop a net baryon number of this magnitude. Finally, the

constancy of the entropy per co-moving volume allows us to characterize the size of co-moving volume corresponding to our present Hubble volume in a very physical way: by the entropy it contains,

$$S_U = \frac{4\pi}{3} H_0^{-3} s \simeq 10^{90}. \quad (16)$$

The standard cosmology is tested back to times as early as about 0.01 s; it is only natural to ask how far back one can sensibly extrapolate. Since the fundamental particles of Nature are point-like quarks and leptons whose interactions are perturbatively weak at energies much greater than 1 GeV, one can imagine extrapolating as far back as the epoch where general relativity becomes suspect, i.e., where quantum gravitational effects are likely to be important: the Planck epoch,  $t \sim 10^{-43}$  s and  $T \sim 10^{19}$  GeV. Of course, at present, our firm understanding of the elementary particles and their interactions only extends to energies of the order of 100 GeV, which corresponds to a time of the order of  $10^{-11}$  s or so. We can be relatively certain that at a temperature of 100–200 MeV ( $t \sim 10^{-5}$  s) there was a transition (likely a second-order phase transition) from quark/gluon plasma to very hot hadronic matter, and that some kind of phase transition associated with the symmetry breakdown of the electroweak theory took place at a temperature of the order of 300 GeV ( $t \sim 10^{-11}$  s).

### 2.3 The concept of particle horizon

In spite of the fact that the universe was vanishingly small at early times, the rapid expansion precluded causal contact from being established throughout. Photons travel on null paths characterized by  $dr = dt/a(t)$ ; the physical distance that a photon could have travelled since the bang until time  $t$ , the distance to the particle horizon, is

$$\begin{aligned} R_H(t) &= a(t) \int_0^t \frac{dt'}{a(t')} \\ &= \frac{t}{(1-n)} = n \frac{H^{-1}}{(1-n)} \sim H^{-1} \quad \text{for } a(t) \propto t^n, \quad n < 1. \end{aligned} \quad (17)$$

Using the conformal time  $d\tau = dt/a$ , the particle horizon becomes

$$R_H(t) = a(\tau) \int_{\tau_0}^{\tau} d\tau, \quad (18)$$

where  $\tau_0$  indicates the conformal time corresponding to  $t = 0$ . Note, in the standard cosmology the distance to the horizon is finite, and up to numerical factors, equal to the age of the universe or the Hubble radius,  $H^{-1}$ . For this reason, we shall use horizon and Hubble radius interchangeably<sup>1</sup>.

Note also that a physical length scale  $\lambda$  is within the horizon if  $\lambda < R_H \sim H^{-1}$ . Since we can identify the length scale  $\lambda$  with its wavenumber  $k$ ,  $\lambda = 2\pi a/k$ , we shall have the following rule

$$\begin{aligned} \frac{k}{aH} &\ll 1 \implies \text{SCALE } \lambda \text{ OUTSIDE THE HORIZON} \\ \frac{k}{aH} &\gg 1 \implies \text{SCALE } \lambda \text{ WITHIN THE HORIZON} \end{aligned}$$

<sup>1</sup>As we shall see, in inflationary models the horizon and Hubble radius are not roughly equal as the horizon distance grows exponentially relative to the Hubble radius; in fact, at the end of inflation they differ by  $e^N$ , where  $N$  is the number of e-folds of inflation. However, we shall slip and use “horizon” and “Hubble radius” interchangeably, though we shall always mean Hubble radius.

### 3 The shortcomings of the standard Big Bang theory

By now the shortcomings of standard cosmology are well appreciated: the horizon or large-scale smoothness problem; the small-scale inhomogeneity problem (origin of density perturbations); and the flatness or oldness problem. we shall briefly review only the horizon problem here here.

#### 3.1 The horizon problem

According to standard cosmology, photons decoupled from the rest of the components (electrons and baryons) at a temperature of the order of 0.3 eV. This corresponds to the so-called surface of ‘last scattering’ at a redshift of about 1100 and an age of about  $180000 (\Omega_0 h^2)^{-1/2}$  yr. From the epoch of last scattering onwards, photons free-stream and reach us basically untouched. Detecting primordial photons is therefore equivalent to take a picture of the universe when the latter was about 300 000 years old. The spectrum of the cosmic background radiation (CBR) is consistent with that of a black body at temperature 2.73 K over more than three decades in wavelength.

The most accurate measurement of the temperature and spectrum is that by the WMAP5 instrument on the COBE satellite which determined its temperature to be  $2.726 \pm 0.01$  K [9]. The length corresponding to our present Hubble radius (which is approximately the radius of our observable universe) at the time of last scattering was

$$\lambda_H(t_{LS}) = R_H(t_0) \left( \frac{a_{LS}}{a_0} \right) = R_H(t_0) \left( \frac{T_0}{T_{LS}} \right).$$

On the other hand, during the matter-dominated period, the Hubble length decreased with a different law

$$H^2 \propto \rho_M \propto a^{-3} \propto T^3.$$

At last-scattering

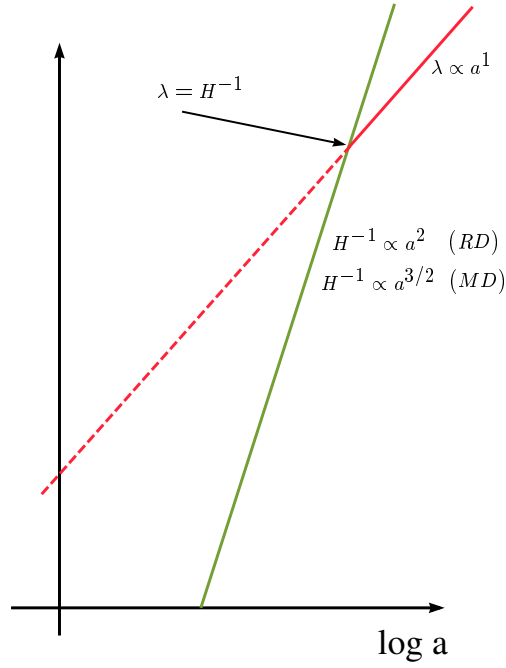
$$H_{LS}^{-1} = R_H(t_0) \left( \frac{T_{LS}}{T_0} \right)^{-3/2} \ll R_H(t_0).$$

The length corresponding to our present Hubble radius was much larger than the horizon at that time. This can be shown by comparing the volumes corresponding to these two scales

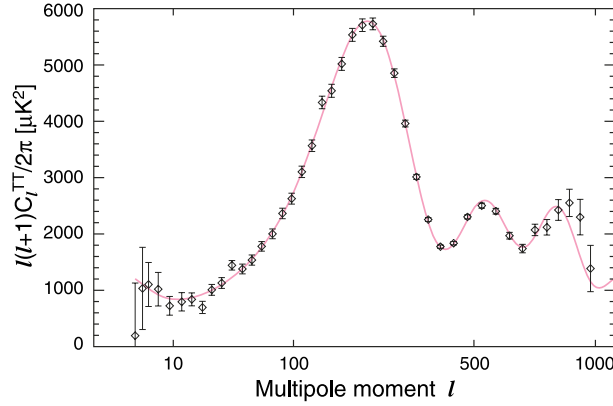
$$\frac{\lambda_H^3(T_{LS})}{H_{LS}^{-3}} = \left( \frac{T_0}{T_{LS}} \right)^{-3/2} \approx 10^6. \quad (19)$$

There were  $\sim 10^6$  causally disconnected regions within the volume that now corresponds to our horizon! It is difficult to come up with a process other than an early hot and dense phase in the history of the universe that would lead to a precise black body for a bath of photons which were causally disconnected the last time they interacted with the surrounding plasma.

The horizon problem is well represented by Fig. 1 where the solid line indicates the horizon scale and the dashed line any generic physical length scale  $\lambda$ . Suppose, indeed, that  $\lambda$  indicates the distance between two photons we detect today. From Eq. (19) we discover that at the time of emission (last-scattering) the two photons could not talk to each other, the dashed line is above the solid line. There is another aspect of the horizon problem which is related to the problem of initial conditions for the cosmological perturbations. We have every indication that the universe at early times, say  $t \ll 300\,000$  yr, was very homogeneous; however, today inhomogeneity (or structure) is ubiquitous: stars ( $\delta\rho/\rho \sim 10^{30}$ ), galaxies ( $\delta\rho/\rho \sim 10^5$ ), clusters of galaxies ( $\delta\rho/\rho \sim 10\text{--}10^3$ ), superclusters, or “clusters of clusters” ( $\delta\rho/\rho \sim 1$ ), voids ( $\delta\rho/\rho \sim -1$ ), great walls, and so on. For some twenty-five years standard cosmology has provided a general framework for understanding this picture. Once the universe becomes matter dominated (around 1000 yr after the bang) primeval density inhomogeneities ( $\delta\rho/\rho \sim 10^{-5}$ ) are amplified by gravity and grow into the structure we see today [2]. The existence of density inhomogeneities



**Fig. 1:** The horizon scale (solid line) and a physical scale  $\lambda$  (dashed line) as function of the scale factor  $a$



**Fig. 2:** The CMBR anisotropy as function of  $\ell$  (from Ref. [9])

has another important consequence: fluctuations in the temperature of the CMB radiation of a similar amplitude. The temperature difference measured between two points separated by a large angle ( $\gtrsim 1^\circ$ ) arises due to a very simple physical effect: the difference in the gravitational potential between the two points on the last scattering surface, which in turn is related to the density perturbation, determines the temperature anisotropy on the angular scale subtended by that length scale,

$$\left(\frac{\delta T}{T}\right)_\theta \approx \left(\frac{\delta \rho}{\rho}\right)_\lambda, \quad (20)$$

where the scale  $\lambda \sim 100h^{-1} \text{ Mpc}(\theta/\text{deg})$  subtends an angle  $\theta$  on the last-scattering surface. This is known as the Sachs–Wolfe effect [11, 12]. We shall come back to this piece of physics. The temperature

anisotropy is commonly expanded in spherical harmonics

$$\frac{\Delta T}{T}(x_0, \tau_0, \mathbf{n}) = \sum_{\ell m} a_{\ell, m}(x_0) Y_{\ell m}(\mathbf{n}), \quad (21)$$

where  $x_0$  and  $\tau_0$  are our position and the preset time, respectively,  $\mathbf{n}$  is the direction of observation,  $\ell$ 's are the different multipoles and<sup>2</sup>

$$\langle a_{\ell m} a_{\ell' m'}^* \rangle = \delta_{\ell, \ell'} \delta_{m, m'} C_\ell, \quad (22)$$

where the deltas are due to the fact that the process that created the anisotropy is statistically isotropic. The  $C_\ell$ 's are the so-called CMB power spectrum. For homogeneity and isotropy, the  $C_\ell$ 's are neither a function of  $x_0$ , nor of  $m$ . The two-point correlation function is related to the  $C_\ell$ 's in the following way

$$\begin{aligned} \left\langle \frac{\delta T(\mathbf{n})}{T} \frac{\delta T(\mathbf{n}')}{T} \right\rangle &= \sum_{\ell \ell' m m'} \langle a_{\ell m} a_{\ell' m'}^* \rangle Y_{\ell m}(\mathbf{n}) Y_{\ell' m'}^*(\mathbf{n}') \\ &= \sum_{\ell} C_\ell \sum_m Y_{\ell m}(\mathbf{n}) Y_{\ell m}^*(\mathbf{n}') = \frac{1}{4\pi} \sum_{\ell} (2\ell + 1) C_\ell P_\ell(\mu = \mathbf{n} \cdot \mathbf{n}') \end{aligned} \quad (23)$$

where we have used the addition theorem for the spherical harmonics, and  $P_\ell$  is the Legendre polynomial of order  $\ell$ . In expression (23) the expectation value is an ensemble average. It can be regarded as an average over the possible observer positions, but not in general as an average over the single sky we observe, because of the cosmic variance<sup>3</sup>. WMAP5 data are given in Fig. 2.

Let us now consider the last scattering surface. In co-moving coordinates the latter is ‘far’ from us a distance equal to

$$\int_{t_{\text{LS}}}^{t_0} \frac{dt}{a} = \int_{\tau_{\text{LS}}}^{\tau_0} d\tau = (\tau_0 - \tau_{\text{LS}}). \quad (24)$$

A given co-moving scale  $\lambda$  is therefore projected on the last scattering surface sky on an angular scale

$$\theta \simeq \frac{\lambda}{(\tau_0 - \tau_{\text{LS}})}, \quad (25)$$

where we have neglected tiny curvature effects. Consider now that the scale  $\lambda$  is of the order of the co-moving sound horizon at the time of last scattering,  $\lambda \sim c_s \tau_{\text{LS}}$ , where  $c_s \simeq 1/\sqrt{3}$  is the sound velocity at which photons propagate in the plasma at the last scattering. This corresponds to an angle

$$\theta \simeq c_s \frac{\tau_{\text{LS}}}{(\tau_0 - \tau_{\text{LS}})} \simeq c_s \frac{\tau_{\text{LS}}}{\tau_0}, \quad (26)$$

where the last passage has been performed knowing that  $\tau_0 \gg \tau_{\text{LS}}$ . Since the universe is matter-dominated from the time of last scattering onwards, the scale factor has the following behaviour:  $a \sim T^{-1} \sim t^{2/3} \sim \tau^2$ . The angle  $\theta_{\text{HOR}}$  subtended by the sound horizon on the last scattering surface then becomes

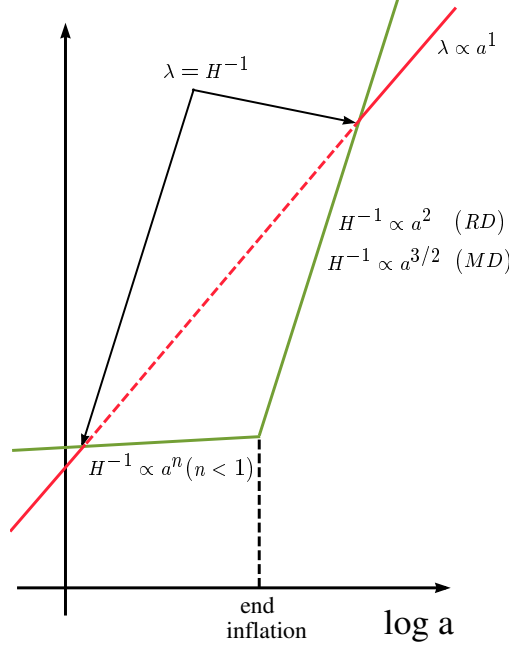
$$\theta_{\text{HOR}} \simeq c_s \left( \frac{T_0}{T_{\text{LS}}} \right)^{1/2} \sim 1^\circ, \quad (27)$$

where we have used  $T_{\text{LS}} \simeq 0.3 \text{ eV}$  and  $T_0 \sim 10^{-13} \text{ GeV}$ . This corresponds to a multipole  $\ell_{\text{HOR}}$

<sup>2</sup>An alternative definition is  $C_\ell = \langle |a_{\ell m}|^2 \rangle = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2$ .

<sup>3</sup>The usual hypothesis is that we observe a typical realization of the ensemble. This means that we expect the difference between the observed values  $|a_{\ell m}|^2$  and the ensemble averages  $C_\ell$  to be of the order of the mean-square deviation of  $|a_{\ell m}|^2$  from  $C_\ell$ . The latter is called cosmic variance and, because we are dealing with a Gaussian distribution, it is equal to  $2C_\ell$  for each multipole  $\ell$ . For a single  $\ell$ , averaging over the  $(2\ell + 1)$  values of  $m$  reduces the cosmic variance by a factor  $(2\ell + 1)$ , but it remains a serious limitation for low multipoles.





**Fig. 3:** The behaviour of a generic scale  $\lambda$  and the horizon scale  $H^{-1}$  in the standard inflationary model

$$\ell_{\text{HOR}} = \frac{\pi}{\theta_{\text{HOR}}} \simeq 200. \quad (28)$$

From these estimates we conclude that two photons which on the last scattering surface were separated by an angle larger than  $\theta_{\text{HOR}}$ , corresponding to multipoles smaller than  $\ell_{\text{HOR}} \sim 200$ , were not in causal contact. On the other hand, from Fig. 2 it is clear that small anisotropies, of the *same* order of magnitude  $\delta T/T \sim 10^{-5}$  are present at  $\ell \ll 200$ . We conclude that one of the striking features of the CMB fluctuations is that they appear to be non-causal. Photons at the last scattering surface which were causally disconnected have the same small anisotropies! The existence of particle horizons in the standard cosmology precludes explaining the smoothness as a result of microphysical events: the horizon at decoupling, the last time one could imagine temperature fluctuations being smoothed by particle interactions, corresponds to an angular scale on the sky of about  $1^\circ$ , which precludes temperature variations on larger scales from being erased. To account for the small-scale lumpiness of the universe today, density perturbations with horizon-crossing amplitudes of  $10^{-5}$  on scales of 1 Mpc to  $10^4$  Mpc or so are required.

As can be seen in Fig. 1, in the standard cosmology the physical size of a perturbation, which grows as the scale factor, begins larger than the horizon and, relatively late in the history of the universe, crosses inside the horizon. This precludes a causal microphysical explanation for the origin of the required density perturbations.

From the considerations made so far, it appears that solving the horizon problem of the standard Big Bang theory requires that the universe go through a primordial period during which the physical scales  $\lambda$  evolve faster than the horizon scale  $H^{-1}$ .

If there is period during which physical length scales grow faster than  $H^{-1}$ , length scales  $\lambda$  which are within the horizon today,  $\lambda < H^{-1}$  (such as the distance between two detected photons) and were outside the horizon for some period,  $\lambda > H^{-1}$  (for instance at the time of last scattering when the two

photons were emitted), had a chance to be within the horizon at some primordial epoch,  $\lambda < H^{-1}$  again, see Fig. 3. If this happens, the homogeneity and the isotropy of the CMB can easily be explained: photons that we receive today and were emitted from the last scattering surface from causally disconnected regions have the same temperature because they had a chance to ‘talk’ to each other at some primordial stage of the evolution of the universe.

The second condition can easily be expressed as a condition on the scale factor  $a$ . Since a given scale  $\lambda$  scales like  $\lambda \sim a$  and  $H^{-1} = a/\dot{a}$ , we need to impose that there is a period during which

$$\left( \frac{\lambda}{H^{-1}} \right)' = \ddot{a} > 0.$$

We can therefore introduce the following rigorous definition: an inflationary stage is a period of the universe during which the latter accelerates

$$\text{INFLATION} \iff \ddot{a} > 0.$$

*Comment:* Let us stress that during such an accelerating phase the universe expands *adiabatically*. This means that during inflation one can exploit the usual FRW equations (3) and (5). It must be clear therefore that the non-adiabaticity condition is satisfied not during inflation, but during the phase transition between the end of inflation and the beginning of the radiation-dominated phase. At this transition phase a large entropy is generated under the form of relativistic degrees of freedom: the Big Bang has taken place.

#### 4 The standard inflationary universe

From the previous section we have learned that an accelerating stage during the primordial phases of the evolution of the universe might be able to solve the horizon problem. From Eq. (5) we learn that

$$\ddot{a} > 0 \iff (\rho + 3P) < 0.$$

An accelerating period is obtainable only if the overall pressure  $p$  of the universe is negative:  $p < -\rho/3$ . Neither a radiation-dominated phase nor a matter-dominated phase (for which  $p = \rho/3$  and  $p = 0$ , respectively) satisfy such a condition. Let us postpone for the time being the problem of finding a ‘candidate’ able to provide the condition  $P < -\rho/3$ . For sure, inflation is a phase of the history of the universe occurring before the era of nucleosynthesis ( $t \approx 1$  s,  $T \approx 1$  MeV) during which the light elements abundances were formed. This is because nucleosynthesis is the earliest epoch from which we have experimental data and they are in agreement with the predictions of the standard Big Bang theory. However, the thermal history of the universe before the epoch of nucleosynthesis is unknown.

In order to study the properties of the period of inflation, we assume the extreme condition  $p = -\rho$  which considerably simplifies the analysis. A period of the universe during which  $P = -\rho$  is called the de Sitter stage. By inspecting Eqs. (3) and (4), we learn that during the de Sitter phase

$$\begin{aligned} \rho &= \text{constant}, \\ H_I &= \text{constant}, \end{aligned}$$

where we have indicated by  $H_I$  the value of the Hubble rate during inflation. Correspondingly, solving Eq. (3) gives

$$a = a_i e^{H_I(t-t_i)}, \tag{29}$$

where  $t_i$  denotes the time at which inflation starts. Let us now see how such a period of exponential expansion takes care of the shortcomings of the standard Big Bang Theory<sup>4</sup>.

#### 4.1 Inflation and the horizon problem

During the inflationary (de Sitter) epoch the horizon scale  $H^{-1}$  is constant. If inflation lasts long enough, all the physical scales that have left the horizon during the radiation-dominated or matter-dominated phase can re-enter the horizon in the past: this is because such scales are exponentially reduced. As we have seen in the previous section, this explains both the problem of the homogeneity of CMB and the initial condition problem of small cosmological perturbations. Once the physical length is within the horizon, microphysics can act, the universe can be made approximately homogeneous and the primeval inhomogeneities can be created.

Let us see how long inflation must be sustained in order to solve the horizon problem. Let  $t_i$  and  $t_f$  be, respectively, the time of beginning and end of inflation. We can define the corresponding number of e-foldings  $N$

$$N = \ln [H_I(t_e - t_i)]. \quad (30)$$

A necessary condition to solve the horizon problem is that the largest scale we observe today, the present horizon  $H_0^{-1}$ , was reduced during inflation to a value  $\lambda_{H_0}(t_i)$  smaller than the value of horizon length  $H_I^{-1}$  during inflation. This gives

$$\lambda_{H_0}(t_i) = H_0^{-1} \left( \frac{a_{t_f}}{a_{t_0}} \right) \left( \frac{a_{t_i}}{a_{t_f}} \right) = H_0^{-1} \left( \frac{T_0}{T_f} \right) e^{-N} \lesssim H_I^{-1},$$

where we have neglected for simplicity the short period of matter-domination and we have called  $T_f$  the temperature at the end of inflation (to be identified with the reheating temperature  $T_{RH}$  at the beginning of the radiation-dominated phase after inflation, see later). We get

$$N \gtrsim \ln \left( \frac{T_0}{H_0} \right) - \ln \left( \frac{T_f}{H_I} \right) \approx 67 + \ln \left( \frac{T_f}{H_I} \right).$$

Apart from the logarithmic dependence, we obtain  $N \gtrsim 70$ .

#### 4.2 A prediction of inflation

Since during inflation the Hubble rate is constant

$$\Omega - 1 = \frac{k}{a^2 H^2} \propto \frac{1}{a^2}.$$

On the other hand it is easy to show that to reproduce a value of  $(\Omega_0 - 1)$  of order of unity today, the initial value of  $(\Omega - 1)$  at the beginning of the radiation-dominated phase must be  $|\Omega - 1| \sim 10^{-60}$ . Since we identify the beginning of the radiation-dominated phase with the beginning of inflation, we require

$$|\Omega - 1|_{t=t_f} \sim 10^{-60}.$$

During inflation

$$\frac{|\Omega - 1|_{t=t_f}}{|\Omega - 1|_{t=t_i}} = \left( \frac{a_i}{a_f} \right)^2 = e^{-2N}. \quad (31)$$

Taking  $|\Omega - 1|_{t=t_i}$  of order unity, it is enough to require that  $N \approx 70$ . However, IF the period of inflation lasts longer than 70 e-foldings the present-day value of  $\Omega_0$  will be equal to unity with great precision. One can say that a generic prediction of inflation is that

---

<sup>4</sup>Despite the fact that the growth of the scale factor is exponential and the expansion is *superluminal*, this is not in contradiction with what is dictated by relativity. Indeed, it is the spacetime itself which is propagating so fast and not a light signal in it.

$$\text{INFLATION} \implies \Omega_0 = 1.$$

This statement, however, must be taken *cum grano salis* and properly specified. Inflation does not change the global geometric properties of the space-time. If the universe is open or closed, it will always remain flat or closed, independently from inflation. What inflation does is to magnify the radius of curvature  $R_{\text{curv}}$  defined in Eq. (9) so that locally the universe is flat with a great precision. As we shall see, the current data on the CMB anisotropies confirm this prediction.

### 4.3 Inflation and the inflaton

In the previous subsections we have described the various advantages of having a period of accelerating phase. The latter required  $P < -\rho/3$ . Now, we would like to show that this condition can be attained by means of a simple scalar field. We shall call this field the *inflaton*  $\phi$ .

The action of the inflaton field reads

$$S = \int d^4x \sqrt{-g} \mathcal{L} = \int d^4x \sqrt{-g} \left[ \frac{1}{2} \partial_\mu \phi \partial^\mu \phi + V(\phi) \right], \quad (32)$$

where  $\sqrt{-g} = a^3$  for the FRW metric (1). From the Euler–Lagrange equations

$$\partial^\mu \frac{\delta(\sqrt{-g} \mathcal{L})}{\delta \partial^\mu \phi} - \frac{\delta(\sqrt{-g} \mathcal{L})}{\delta \phi} = 0, \quad (33)$$

we obtain

$$\ddot{\phi} + 3H\dot{\phi} - \frac{\nabla^2 \phi}{a^2} + V'(\phi) = 0, \quad (34)$$

where  $V'(\phi) = (dV(\phi)/d\phi)$ . Note, in particular, the appearance of the friction term  $3H\dot{\phi}$ : a scalar field rolling down its potential suffers a friction due to the expansion of the universe.

We can write the energy momentum tensor of the scalar field

$$T_{\mu\nu} = \partial_\mu \phi \partial_\nu \phi - g_{\mu\nu} \mathcal{L}.$$

The corresponding energy density  $\rho_\phi$  and pressure density  $P_\phi$  are

$$T_{00} = \rho_\phi = \frac{\dot{\phi}^2}{2} + V(\phi) + \frac{(\nabla \phi)^2}{2a^2}, \quad (35)$$

$$T_{ii} = P_\phi = \frac{\dot{\phi}^2}{2} - V(\phi) - \frac{(\nabla \phi)^2}{6a^2}. \quad (36)$$

Note that, if the gradient term were dominant, we would obtain  $P_\phi = -\frac{\rho_\phi}{3}$ , not enough to drive inflation. We can now split the inflaton field in

$$\phi(t) = \phi_0(t) + \delta\phi(\mathbf{x}, t),$$

where  $\phi_0$  is the ‘classical’ (infinite wavelength) field, that is the expectation value of the inflaton field on the initial isotropic and homogeneous state, while  $\delta\phi(\mathbf{x}, t)$  represents the quantum fluctuations around  $\phi_0$ . In this section, we shall be concerned only with the evolution of the classical field  $\phi_0$ . The next section will be devoted to the crucial issue of the evolution of quantum perturbations during inflation. This separation is justified by the fact that quantum fluctuations are much smaller than the classical value and therefore negligible when looking at the classical evolution. Not to be overwhelmed by the notation,

we shall indicate the classical value of the inflaton field by  $\phi$  from now on. The energy momentum tensor becomes

$$T_{00} = \rho_\phi = \frac{\dot{\phi}^2}{2} + V(\phi) \quad (37)$$

$$T_{ii} = P_\phi = \frac{\dot{\phi}^2}{2} - V(\phi). \quad (38)$$

If

$$V(\phi) \gg \dot{\phi}^2$$

we obtain the following condition

$$P_\phi \simeq -\rho_\phi.$$

From this simple calculation, we realize that a scalar field whose energy is dominant in the universe and whose potential energy dominates over the kinetic term gives inflation. Inflation is driven by the vacuum energy of the inflaton field.

#### 4.4 Slow-roll conditions

Let us now quantify better under which circumstances a scalar field may give rise to a period of inflation. The equation of motion of the field is

$$\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0. \quad (39)$$

If we require that  $\dot{\phi}^2 \ll V(\phi)$ , the scalar field is slowly rolling down its potential. This is the reason why such a period is called *slow-roll*. We may also expect that since the potential is flat,  $\ddot{\phi}$  is negligible as well. We shall assume that this is true and we shall quantify this condition soon. The FRW equation (3) becomes

$$H^2 \simeq \frac{8\pi G}{3} V(\phi), \quad (40)$$

where we have assumed that the inflaton field dominates the energy density of the universe. The new equation of motion becomes

$$3H\dot{\phi} = -V'(\phi) \quad (41)$$

which gives  $\dot{\phi}$  as a function of  $V'(\phi)$ . Using Eq. (41) slow-roll conditions then require

$$\dot{\phi}^2 \ll V(\phi) \implies \frac{(V')^2}{V} \ll H^2$$

and

$$\ddot{\phi} \ll 3H\dot{\phi} \implies V'' \ll H^2.$$

It is now useful to define the slow-roll parameters  $\epsilon$  and  $\eta$  in the following way

$$\begin{aligned} \epsilon &= -\frac{\dot{H}}{H^2} = 4\pi G \frac{\dot{\phi}^2}{H^2} = \frac{1}{16\pi G} \left( \frac{V'}{V} \right)^2, \\ \eta &= \frac{1}{8\pi G} \left( \frac{V''}{V} \right) = \frac{1}{3} \frac{V''}{H^2}, \\ \delta &= \eta - \epsilon = -\frac{\ddot{\phi}}{H\dot{\phi}}. \end{aligned}$$

It might be useful to have the same parameters expressed in terms of conformal time

$$\begin{aligned}\epsilon &= 1 - \frac{\mathcal{H}'}{\mathcal{H}^2} = 4\pi G \frac{\phi'^2}{\mathcal{H}^2} \\ \delta &= \eta - \epsilon = 1 - \frac{\phi''}{\mathcal{H}\phi'}.\end{aligned}$$

The parameter  $\epsilon$  quantifies how much the Hubble rate  $H$  changes with time during inflation. Notice that, since

$$\frac{\ddot{a}}{a} = \dot{H} + H^2 = (1 - \epsilon) H^2,$$

inflation can be attained only if  $\epsilon < 1$ :

$$\text{INFLATION} \iff \epsilon < 1.$$

As soon as this condition fails, inflation ends. In general, slow-roll inflation is attained if  $\epsilon \ll 1$  and  $|\eta| \ll 1$ . During inflation the slow-roll parameters  $\epsilon$  and  $\eta$  can be considered to be approximately constant since the potential  $V(\phi)$  is very flat.

*Comment:* In the following, we shall work at *first-order* perturbation in the slow-roll parameters, that is we shall take only the first power of them. Since, using their definition, it is easy to see that  $\dot{\epsilon}, \dot{\eta} = \mathcal{O}(\epsilon^2, \eta^2)$ , this amounts to saying that we shall treat the slow-roll parameters as constant in time.

Within these approximations, it is easy to compute the number of e-foldings between the beginning and the end of inflation. If we indicate by  $\phi_i$  and  $\phi_f$  the values of the inflaton field at the beginning and at the end of inflation, respectively, we find that the *total* number of e-foldings is

$$\begin{aligned}N &\equiv \int_{t_i}^{t_f} H dt \\ &\simeq H \int_{\phi_i}^{\phi_f} \frac{d\phi}{\dot{\phi}} \\ &\simeq -3H^2 \int_{\phi_i}^{\phi_f} \frac{d\phi}{V'} \\ &\simeq -8\pi G \int_{\phi_i}^{\phi_f} \frac{V}{V'} d\phi.\end{aligned}\tag{42}$$

We may also compute the number of e-foldings  $\Delta N$  which are left to go to the end of inflation

$$\Delta N \simeq 8\pi G \int_{\phi_f}^{\phi_{\Delta N}} \frac{V}{V'} d\phi,\tag{43}$$

where  $\phi_{\Delta N}$  is the value of the inflaton field when there are  $\Delta N$  e-foldings to the end of inflation.

1. *Comment:* According to the criterion given in Subsection 2.4, a given scale length  $\lambda = a/k$  leaves the horizon when  $k = aH_k$  where  $H_k$  is the value of the Hubble rate at that time. One can easily compute the rate of change of  $H_k^2$  as a function of  $k$

$$\frac{d \ln H_k^2}{d \ln k} = \left( \frac{d \ln H_k^2}{dt} \right) \left( \frac{dt}{d \ln a} \right) \left( \frac{d \ln a}{d \ln k} \right) = 2 \frac{\dot{H}}{H} \times \frac{1}{H} \times 1 = 2 \frac{\dot{H}}{H^2} = -2\epsilon.\tag{44}$$

2. *Comment:* Take a given physical scale  $\lambda$  today which crossed the horizon scale during inflation. This happened when

$$\lambda \left( \frac{a_f}{a_0} \right) e^{-\Delta N_\lambda} = \lambda \left( \frac{T_0}{T_f} \right) e^{-\Delta N_\lambda} = H_I^{-1}$$

where  $\Delta N_\lambda$  indicates the number of e-foldings from the time the scale crossed the horizon during inflation and the end of inflation. This relation gives a way to determine the number of e-foldings to the end of inflation corresponding to a given scale

$$\Delta N_\lambda \simeq 65 + \ln \left( \frac{\lambda}{3000 \text{ Mpc}} \right) + 2 \ln \left( \frac{V^{1/4}}{10^{14} \text{ GeV}} \right) - \ln \left( \frac{T_f}{10^{10} \text{ GeV}} \right).$$

Scales relevant for the CMB anisotropies correspond to  $\Delta N \sim 60$ .

Inflation ended when the potential energy associated with the inflaton field became smaller than the kinetic energy of the field. By that time, any pre-inflation entropy in the universe had been inflated away, and the energy of the universe was entirely in the form of coherent oscillations of the inflaton condensate around the minimum of its potential. The universe may be said to be frozen after the end of inflation. We know that somehow the low-entropy cold universe dominated by the energy of coherent motion of the  $\phi$  field must be transformed into a high-entropy hot universe dominated by radiation. The process by which the energy of the inflaton field is transferred from the inflaton field to radiation has been dubbed *reheating*. In the theory of reheating, the simplest way to envisage this process is if the co-moving energy density in the zero mode of the inflaton decays into normal particles, which then scatter and thermalize to form a thermal background. It is usually assumed that the decay width of this process is the same as the decay width of a free inflaton field.

Of particular interest is a quantity usually known as the reheat temperature, denoted as  $T_{RH}$ <sup>5</sup>. The reheat temperature is calculated by assuming an instantaneous conversion of the energy density in the inflaton field into radiation when the decay width of the inflaton energy,  $\Gamma_\phi$ , is equal to  $H$ , the expansion rate of the universe.

The reheat temperature is calculated quite easily. After inflation the inflaton field executes coherent oscillations about the minimum of the potential. Averaged over several oscillations, the coherent oscillation energy density redshifts as matter:  $\rho_\phi \propto a^{-3}$ , where  $a$  is the Robertson–Walker scale factor. If we denote as  $\rho_I$  and  $a_I$  the total inflaton energy density and the scale factor at the initiation of coherent oscillations, then the Hubble expansion rate as a function of  $a$  is

$$H^2(a) = \frac{8\pi}{3} \frac{\rho_I}{m_{\text{Pl}}^2} \left( \frac{a_I}{a} \right)^3. \quad (45)$$

Equating  $H(a)$  and  $\Gamma_\phi$  leads to an expression for  $a_I/a$ . Now if we assume that all available coherent energy density is instantaneously converted into radiation at this value of  $a_I/a$ , we can find the reheat temperature by setting the coherent energy density,  $\rho_\phi = \rho_I(a_I/a)^3$ , equal to the radiation energy density,  $\rho_R = (\pi^2/30)g_*T_{RH}^4$ , where  $g_*$  is the effective number of relativistic degrees of freedom at temperature  $T_{RH}$ . The result is

$$T_{RH} = \left( \frac{90}{8\pi^3 g_*} \right)^{1/4} \sqrt{\Gamma_\phi m_{\text{Pl}}} = 0.2 \left( \frac{200}{g_*} \right)^{1/4} \sqrt{\Gamma_\phi m_{\text{Pl}}}. \quad (46)$$

## 5 Inflation and the cosmological perturbations

As we have seen in the previous section, the early universe was made very nearly uniform by a primordial inflationary stage. However, the important caveat in that statement is the word ‘nearly’. Our current understanding of the origin of structure in the universe is that it originated from small ‘seed’ perturbations,

---

<sup>5</sup>So far, we have indicated it by  $T_f$ .

which over time grew to become all of the structure we observe. Once the universe becomes matter dominated (around 1000 yrs after the bang) primeval density inhomogeneities ( $\delta\rho/\rho \sim 10^{-5}$ ) are amplified by gravity and grow into the structure we see today [4]. The fact that a fluid of self-gravitating particles is unstable to the growth of small inhomogeneities was first pointed out by Jeans and is known as the Jeans instability. Furthermore, the existence of these inhomogeneities was confirmed by the COBE discovery of CMB anisotropies; the temperature anisotropies detected almost certainly owe their existence to primeval density inhomogeneities, since, as we have seen, causality precludes microphysical processes from producing anisotropies on angular scales larger than about  $1^\circ$ , the angular size of the horizon at last-scattering.

The growth of small matter inhomogeneities of wavelength smaller than the Hubble scale ( $\lambda \lesssim H^{-1}$ ) is governed by a Newtonian equation:

$$\ddot{\delta}_{\mathbf{k}} + 2H\dot{\delta}_{\mathbf{k}} + v_s^2 \frac{k^2}{a^2} \delta_{\mathbf{k}} = 4\pi G\rho_M \delta_{\mathbf{k}}, \quad (47)$$

where  $v_s^2 = \partial P/\partial\rho_M$  is the square of the speed of sound and we have expanded the perturbation to the matter density in plane waves

$$\frac{\delta\rho_m(\mathbf{x}, t)}{\rho_m} = \frac{1}{(2\pi)^3} \int d^3k \delta_{\mathbf{k}}(t) e^{-i\mathbf{k}\cdot\mathbf{x}}. \quad (48)$$

Competition between the pressure term and the gravity term on the rhs of Eq. (47) determines whether or not pressure can counteract gravity: perturbations with wavenumber larger than the Jeans wavenumber,  $k_J^2 = 4\pi G a^2 \rho_m / v_s^2$ , are Jeans stable and just oscillate; perturbations with smaller wavenumber are Jeans unstable and can grow.

Let us discuss solutions to this equation under different circumstances. First, consider the Jeans problem, evolution of perturbations in a static fluid, i.e.,  $H = 0$ . In this case Jeans unstable perturbations grow exponentially,  $\delta_{\mathbf{k}} \propto \exp(t/\tau)$  where  $\tau = 1/\sqrt{4\pi G\rho_M}$ . Next, consider the growth of Jeans unstable perturbations in a matter-dominated universe, i.e.,  $H^2 = 8\pi G\rho_M/3$  and  $a \propto t^{2/3}$ . Because the expansion tends to pull particles away from one another, the growth is only power law,  $\delta_{\mathbf{k}} \propto t^{2/3}$ ; i.e., at the same rate as the scale factor. Finally, consider a radiation-dominated universe. In this case, the expansion is so rapid that matter perturbations grow very slowly, as  $\ln a$  in a radiation-dominated epoch. Therefore, perturbations may grow only in a matter-dominated period. Once a perturbation reaches an overdensity of order unity or larger it separates from the expansion, i.e., it becomes its own self-gravitating system and ceases to expand any further. In the process of virial relaxation, its size decreases by a factor of two—density increases by a factor of 8; thereafter, its density contrast grows as  $a^3$  since the average matter density is decreasing as  $a^{-3}$ , though smaller scales could become Jeans unstable and collapse further to form smaller objects of higher density.

In order for structure formation to occur via gravitational instability, there must have been small pre-existing fluctuations on physical length scales when they crossed the Hubble radius in the radiation-dominated and matter-dominated eras. In the standard Big Bang model these small perturbations have to be put in by hand, because it is impossible to produce fluctuations on any length scale while it is larger than the horizon. Since the goal of cosmology is to understand the universe on the basis of physical laws, this appeal to initial conditions is unsatisfactory. The challenge is therefore to give an explanation to the small seed perturbations which allow the gravitational growth of the matter perturbations.

Our best guess for the origin of these perturbations is quantum fluctuations during an inflationary era in the early universe. Although originally introduced as a possible solution to the cosmological conundrums such as the horizon, flatness and entropy problems, by far the most useful property of inflation is that it generates spectra of both density perturbations and gravitational waves. These perturbations extend from extremely short scales to scales considerably in excess of the size of the observable universe.

During inflation the scale factor grows quasi-exponentially, while the Hubble radius remains almost constant. Consequently the wavelength of a quantum fluctuation— either in the scalar field whose



potential energy drives inflation or in the graviton field—soon exceeds the Hubble radius. The amplitude of the fluctuation therefore becomes ‘frozen in’. This is quantum mechanics in action at macroscopic scales.

According to quantum field theory, empty space is not entirely empty. It is filled with quantum fluctuations of all types of physical fields. The fluctuations can be regarded as waves of physical fields with all possible wavelengths, moving in all possible directions. If the values of these fields, averaged over some macroscopically large time, vanish then the space filled with these fields seems to us empty and can be called the vacuum.

In the exponentially expanding universe the vacuum structure is much more complicated. The wavelengths of all vacuum fluctuations of the inflaton field  $\phi$  grow exponentially in the expanding universe. When the wavelength of any particular fluctuation becomes greater than  $H^{-1}$ , this fluctuation stops propagating, and its amplitude freezes at some non-zero value  $\delta\phi$  because of the large friction term  $3H\dot{\phi}$  the equation of motion of the field  $\phi$ . The amplitude of this fluctuation then remains almost unchanged for a very long time, whereas its wavelength grows exponentially. Therefore, the appearance of such frozen fluctuation is equivalent to the appearance of a classical field  $\delta\phi$  that does not vanish after having averaged over some macroscopic interval of time. Because the vacuum contains fluctuations of all possible wavelengths, inflation leads to the creation of more and more new perturbations of the classical field with wavelength larger than the horizon scale.

Once inflation has ended, however, the Hubble radius increases faster than the scale factor, so the fluctuations eventually re-enter the Hubble radius during the radiation- or matter-dominated eras. The fluctuations that exit around 60  $e$ -foldings or so before reheating re-enter with physical wavelengths in the range accessible to cosmological observations. These spectra provide a distinctive signature of inflation. They can be measured in a variety of different ways including the analysis of microwave background anisotropies.

Quantum fluctuations of the inflaton field are generated during inflation. Since gravity talks to any component of the universe, small fluctuations of the inflaton field are intimately related to fluctuations of the space-time metric, giving rise to perturbations of the curvature  $\mathcal{R}$  (which will be defined in the following; the reader may loosely think of it as a gravitational potential). The wavelengths  $\lambda$  of these perturbations grow exponentially and leave the horizon soon when  $\lambda > R_H$ . On superhorizon scales, curvature fluctuations are frozen in and may be considered as classical. Finally, when the wavelength of these fluctuations re-enters the horizon, at some radiation- or matter-dominated epoch, the curvature (gravitational potential) perturbations of the space-time give rise to matter (and temperature) perturbations  $\delta\rho$  via the Poisson equation. These fluctuations will then start growing, giving rise to the structures we observe today.

In summary, these are the key ingredients for understanding the observed structures in the universe within the inflationary scenario:

- Quantum fluctuations of the inflaton field are excited during inflation and stretched to cosmological scales. At the same time, being the inflaton fluctuations connected to the metric perturbations through Einstein’s equations, ripples on the metric are also excited and stretched to cosmological scales.
- Gravity acts a messenger since it communicates the small seed perturbations to photons and baryons once a given wavelength becomes smaller than the horizon scale after inflation.

Let us now see how quantum fluctuations are generated during inflation. we shall proceed by steps. First, we shall consider the simplest problem of studying the quantum fluctuations of a generic scalar field during inflation: we shall learn how perturbations evolve as a function of time and compute their spectrum. Then—since a satisfactory description of the generation of quantum fluctuations has to take both the inflaton and the metric perturbations into account— we shall study the system composed

by quantum fluctuations of the inflaton field and quantum fluctuations of the metric.

## 6 Quantum fluctuations of a generic massless scalar field during inflation

Let us first see how the fluctuations of a generic scalar field  $\chi$ , which is *not* the inflaton field, behave during inflation. To warm up we first consider a de Sitter epoch during which the Hubble rate is constant.

### 6.1 Quantum fluctuations of a generic massless scalar field during a de Sitter stage

We assume this field to be massless. The massive case will be analysed in the next subsection.

Expanding the scalar field  $\chi$  in Fourier modes

$$\delta\chi(\mathbf{x}, t) = \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} e^{i\mathbf{k}\cdot\mathbf{x}} \delta\chi_{\mathbf{k}}(t),$$

we can write the equation for the fluctuations as

$$\delta\ddot{\chi}_{\mathbf{k}} + 3H \delta\dot{\chi}_{\mathbf{k}} + \frac{k^2}{a^2} \delta\chi_{\mathbf{k}} = 0. \quad (49)$$

Let us study the qualitative behaviour of the solution to Eq. (49).

- For wavelengths within the horizon,  $\lambda \ll H^{-1}$ , the corresponding wave-number satisfies the relation  $k \gg aH$ . In this regime, we can neglect the friction term  $3H \delta\dot{\chi}_{\mathbf{k}}$  and Eq. (49) reduces to

$$\delta\ddot{\chi}_{\mathbf{k}} + \frac{k^2}{a^2} \delta\chi_{\mathbf{k}} = 0, \quad (50)$$

which is basically the equation of motion of an harmonic oscillator. Of course, the frequency term  $k^2/a^2$  depends upon time because the scale factor  $a$  grows exponentially. On the qualitative level, however, one expects that when the wavelength of the fluctuation is within the horizon, the fluctuation oscillates.

- For wavelengths above the horizon,  $\lambda \gg H^{-1}$ , the corresponding wave-number satisfies the relation  $k \ll aH$  and the term  $k^2/a^2$  can be safely neglected. Equation (49) reduces to

$$\delta\ddot{\chi}_{\mathbf{k}} + 3H \delta\dot{\chi}_{\mathbf{k}} = 0, \quad (51)$$

which tells us that on superhorizon scales  $\delta\chi_{\mathbf{k}}$  remains constant.

We have therefore the following picture: take a given fluctuation whose initial wavelength  $\lambda \sim a/k$  is within the horizon. The fluctuations oscillate till the wavelength becomes of the order of the horizon scale. When the wavelength crosses the horizon, the fluctuation ceases to oscillate and gets frozen in.

Let us now study the evolution of the fluctuation in a more quantitative way. To do so, we perform the following redefinition

$$\delta\chi_{\mathbf{k}} = \frac{\delta\sigma_{\mathbf{k}}}{a}$$

and we work in conformal time  $d\tau = dt/a$ . For the time being, we solve the problem for a pure de Sitter expansion and we take the scale factor exponentially growing as  $a \sim e^{Ht}$ ; the corresponding conformal factor reads (after choosing properly the integration constants)

$$a(\tau) = -\frac{1}{H\tau} \quad (\tau < 0).$$

In the following we shall also solve the problem in the case of quasi de Sitter expansion. The beginning of inflation coincides with some initial time  $\tau_i \ll 0$ . We find that Eq. (49) becomes

$$\delta\sigma_{\mathbf{k}}'' + \left(k^2 - \frac{a''}{a}\right) \delta\sigma_{\mathbf{k}} = 0. \quad (52)$$

We obtain an equation which is very ‘close’ to the equation for a Klein–Gordon scalar field in flat space-time, the only difference being a negative time-dependent mass term  $-a''/a = -2/\tau^2$ . Equation (52) can be obtained from an action of the type

$$\delta S_{\mathbf{k}} = \int d\tau \left[ \frac{1}{2} \delta\sigma_{\mathbf{k}}'^2 - \frac{1}{2} \left(k^2 - \frac{a''}{a}\right) \delta\sigma_{\mathbf{k}}^2 \right], \quad (53)$$

which is the canonical action for a simple harmonic oscillator with canonical commutation relations  $\delta\sigma_{\mathbf{k}}^* \delta\sigma_{\mathbf{k}}' - \delta\sigma_{\mathbf{k}} \delta\sigma_{\mathbf{k}}'^* = -i$ .

Let us study the behaviour of this equation on subhorizon and superhorizon scales. Since

$$\frac{k}{aH} = -k\tau,$$

on subhorizon scales  $k^2 \gg a''/a$  Equation (52) reduces to

$$\delta\sigma_{\mathbf{k}}'' + k^2 \delta\sigma_{\mathbf{k}} = 0,$$

whose solution is a plane wave

$$\delta\sigma_{\mathbf{k}} = \frac{e^{-ik\tau}}{\sqrt{2k}} \quad (k \gg aH). \quad (54)$$

We find again that fluctuations with wavelength within the horizon oscillate exactly like in flat space-time. This does not come as a surprise. In the ultraviolet regime, that is for wavelengths much smaller than the horizon scale, one expects that approximating the space-time as flat is a good approximation.

On superhorizon scales,  $k^2 \ll a''/a$  Equation (52) reduces to

$$\delta\sigma_{\mathbf{k}}'' - \frac{a''}{a} \delta\sigma_{\mathbf{k}} = 0,$$

which is satisfied by

$$\delta\sigma_{\mathbf{k}} = B(k) a \quad (k \ll aH) \quad (55)$$

where  $B(k)$  is a constant of integration. Roughly matching the (absolute values of the) solutions (54) and (55) at  $k = aH$  ( $-k\tau = 1$ ), we can determine the (absolute value of the) constant  $B(k)$

$$|B(k)| a = \frac{1}{\sqrt{2k}} \implies |B(k)| = \frac{1}{a\sqrt{2k}} = \frac{H}{\sqrt{2k^3}}.$$

Going back to the original variable  $\delta\chi_{\mathbf{k}}$ , we obtain that the quantum fluctuation of the  $\chi$  field on superhorizon scales is constant and approximately equal to

$$|\delta\chi_{\mathbf{k}}| \simeq \frac{H}{\sqrt{2k^3}} \quad (\text{ON SUPERHORIZON SCALES})$$

In fact we can do much better, since Eq. (52) has an *exact* solution:

$$\delta\sigma_{\mathbf{k}} = \frac{e^{-ik\tau}}{\sqrt{2k}} \left( 1 + \frac{i}{k\tau} \right). \quad (56)$$

This solution reproduces all that we have found by qualitative arguments in the two extreme regimes  $k \ll aH$  and  $k \gg aH$ . We have performed the matching procedure to show that the latter can be very useful to determine the behaviour of the solution on superhorizon scales when the exact solution is not known.

## 6.2 The power spectrum

Let us define now the power spectrum, a useful quantity to characterize the properties of the perturbations. For a generic quantity  $g(\mathbf{x}, t)$ , which can be expanded in Fourier space as

$$g(\mathbf{x}, t) = \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} e^{i\mathbf{k}\cdot\mathbf{x}} g_{\mathbf{k}}(t),$$

the power spectrum can be defined as

$$\langle 0 | g_{\mathbf{k}_1}^* g_{\mathbf{k}_2} | 0 \rangle \equiv \delta^{(3)}(\mathbf{k}_1 - \mathbf{k}_2) \frac{2\pi^2}{k^3} \mathcal{P}_g(k), \quad (57)$$

where  $|0\rangle$  is the vacuum quantum state of the system. This definition leads to the usual relation

$$\langle 0 | g^2(\mathbf{x}, t) | 0 \rangle = \int \frac{dk}{k} \mathcal{P}_g(k). \quad (58)$$

## 6.3 Quantum fluctuations of a generic scalar field in a quasi de Sitter stage

So far, we have computed the time evolution and the spectrum of the quantum fluctuations of a generic scalar field  $\chi$  supposing that the scale factor evolves like in a pure de Sitter expansion,  $a(\tau) = -1/(H\tau)$ . However, during inflation the Hubble rate is not exactly constant, but changes with time as  $\dot{H} = -\epsilon H^2$  (quasi de Sitter expansion). In this subsection, we shall solve for the perturbations in a quasi de Sitter expansion. Using the definition of the conformal time, one can show that the scale factor for small values of  $\epsilon$  becomes

$$a(\tau) = -\frac{1}{H} \frac{1}{\tau(1-\epsilon)}.$$

The fluctuation mass-squared mass term is

$$M^2(\tau) = m_\chi^2 a^2 - \frac{a''}{a},$$

where

$$\begin{aligned} \frac{a''}{a} &= a^2 \left( \frac{\ddot{a}}{a} + H^2 \right) = a^2 \left( \dot{H} + 2H^2 \right) \\ &= a^2 (2 - \epsilon) H^2 = \frac{(2 - \epsilon)}{\tau^2 (1 - \epsilon)^2} \\ &\simeq \frac{1}{\tau^2} (2 + 3\epsilon). \end{aligned} \quad (59)$$

Armed with these results, we may compute the variance of the perturbations of the generic  $\chi$  field

$$\begin{aligned} \langle 0 | (\delta\chi(\mathbf{x}, t))^2 | 0 \rangle &= \int \frac{d^3k}{(2\pi)^3} |\delta\chi_{\mathbf{k}}|^2 \\ &= \int \frac{dk}{k} \frac{k^3}{2\pi^2} |\delta\chi_{\mathbf{k}}|^2 \\ &= \int \frac{dk}{k} \mathcal{P}_{\delta\chi}(k), \end{aligned} \quad (60)$$

which defines the power spectrum of the fluctuations of the scalar field  $\chi$

$$\mathcal{P}_{\delta\chi}(k) \equiv \frac{k^3}{2\pi^2} |\delta\chi_{\mathbf{k}}|^2. \quad (61)$$

Since we have seen that fluctuations are (nearly) frozen in on superhorizon scales, a way of characterizing the perturbations is to compute the spectrum on scales larger than the horizon. For a massive scalar field, we obtain

$$\mathcal{P}_{\delta\chi}(k) = \left(\frac{H}{2\pi}\right)^2 \left(\frac{k}{aH}\right)^{3-2\nu_\chi}, \quad (62)$$

where, taking  $m_\chi^2/H^2 = 3\eta_\chi$  and expanding for small values of  $\epsilon$  and  $\eta$ ,

$$\nu_\chi \simeq \frac{3}{2} + \epsilon - \eta_\chi. \quad (63)$$

We may also define the *spectral index*  $n_{\delta\chi}$  of the fluctuations as

$$n_{\delta\chi} - 1 = \frac{d\ln \mathcal{P}_{\delta\phi}}{d\ln k} = 3 - 2\nu_\chi = 2\eta_\chi - 2\epsilon.$$

The power spectrum of fluctuations of the scalar field  $\chi$  is therefore *nearly flat*, that is is nearly independent of the wavelength  $\lambda = \pi/k$ : the amplitude of the fluctuation on superhorizon scales does almost not depend upon the time at which the fluctuation crosses the horizon and becomes frozen in. The small tilt of the power spectrum arises from the fact that the scalar field  $\chi$  is massive and because during inflation the Hubble rate is not exactly constant, but nearly constant, where ‘nearly’ is quantified by the slow-roll parameters  $\epsilon$ . Adopting the traditional terminology, we may say that the spectrum of perturbations is blue if  $n_{\delta\chi} > 1$  (more power in the ultraviolet) and red if  $n_{\delta\chi} < 1$  (more power in the infrared). The power spectrum of the perturbations of a generic scalar field  $\chi$  generated during a period of slow-roll inflation may be either blue or red. This depends upon the relative magnitude between  $\eta_\chi$  and  $\epsilon$ .

*Comment:* We might have computed the spectral index of the spectrum  $\mathcal{P}_{\delta\chi}(k)$  by first solving the equation for the perturbations of the field  $\chi$  in a di Sitter stage, with  $H = \text{constant}$  and therefore  $\epsilon = 0$ , and then taking into account the time evolution of the Hubble rate introducing the subscript in  $H_k$  whose time variation is determined by Eq. (44). Correspondingly,  $H_k$  is the value of the Hubble rate when a given wavelength  $\sim k^{-1}$  crosses the horizon (from that point on the fluctuation remains frozen in). The power spectrum in such an approach would read

$$\mathcal{P}_{\delta\chi}(k) = \left(\frac{H_k}{2\pi}\right)^2 \left(\frac{k}{aH}\right)^{3-2\nu_\chi} \quad (64)$$

with  $3 - 2\nu_\chi \simeq \eta_\chi$ . Using Eq. (44), one finds

$$n_{\delta\chi} - 1 = \frac{d\ln \mathcal{P}_{\delta\phi}}{d\ln k} = \frac{d\ln H_k^2}{d\ln k} + 3 - 2\nu_\chi = 2\eta_\chi - 2\epsilon$$

which reproduces our previous findings.

*Comment:* Since on superhorizon scales

$$\delta\chi_{\mathbf{k}} \simeq \frac{H}{\sqrt{2k^3}} \left(\frac{k}{aH}\right)^{\eta_\chi - \epsilon} \simeq \frac{H}{\sqrt{2k^3}} \left[1 + (\eta_\chi - \epsilon) \ln \left(\frac{k}{aH}\right)\right],$$

we discover that

$$|\delta\dot{\chi}_{\mathbf{k}}| \simeq |H(\eta_\chi - \epsilon) \delta\chi_{\mathbf{k}}| \ll |H \delta\chi_{\mathbf{k}}|, \quad (65)$$

that is, on superhorizon scales the time variation of the perturbations can be safely neglected.

## 7 Quantum fluctuations during inflation

As we have mentioned in the previous section, the linear theory of the cosmological perturbations represents a cornerstone of modern cosmology and is used to describe the formation and evolution of structures in the universe as well as the anisotropies of the CMB. The seeds for these inhomogeneities were generated during inflation and stretched over astronomical scales because of the rapid superluminal expansion of the universe during the (quasi) de Sitter epoch.

In the previous section we have already seen that perturbations of a generic scalar field  $\chi$  are generated during a (quasi) de Sitter expansion. The inflaton field is a scalar field and, as such, we conclude that inflaton fluctuations will be generated as well. However, the inflaton is special from the point of view of perturbations. The reason is very simple. By assumption, the inflaton field dominates the energy density of the universe during inflation. Any perturbation in the inflaton field means a perturbation of the stress energy momentum tensor

$$\delta\phi \implies \delta T_{\mu\nu}.$$

A perturbation in the stress energy momentum tensor implies, through Einstein's equations of motion, a perturbation of the metric

$$\delta T_{\mu\nu} \implies \left[ \delta R_{\mu\nu} - \frac{1}{2} \delta (g_{\mu\nu} R) \right] = 8\pi G \delta T_{\mu\nu} \implies \delta g_{\mu\nu}.$$

On the other hand, a perturbation of the metric induces a back-reaction on the evolution of the inflaton perturbation through the perturbed Klein–Gordon equation of the inflaton field

$$\delta g_{\mu\nu} \implies \delta \left( \partial_\mu \partial^\mu \phi + \frac{\partial V}{\partial \phi} \right) = 0 \implies \delta\phi.$$

This logic chain makes us conclude that the perturbations of the inflaton field and of the metric are tightly coupled to each other and have to be studied together

$$\delta\phi \iff \delta g_{\mu\nu}.$$

As we shall see shortly, this relation is stronger than one might think because of the issue of gauge invariance.

Before launching ourselves into the problem of finding the evolution of the quantum perturbations of the inflaton field when they are coupled to gravity, let us give a heuristic explanation of why we expect that during inflation such fluctuations are indeed present.

If we take Eq. (34) and split the inflaton field as its classical value  $\phi_0$  plus the quantum fluctuation  $\delta\phi$ ,  $\phi(\mathbf{x}, t) = \phi_0(t) + \delta\phi(\mathbf{x}, t)$ , the quantum perturbation  $\delta\phi$  satisfies the equation of motion

$$\delta\ddot{\phi} + 3H \delta\dot{\phi} - \frac{\nabla^2 \delta\phi}{a^2} + V'' \delta\phi = 0. \quad (66)$$

Differentiating Eq. (39) wrt time and taking  $H$  constant (de Sitter expansion) we find

$$(\phi_0)''' + 3H\ddot{\phi}_0 + V'' \dot{\phi}_0 = 0. \quad (67)$$

Let us consider for simplicity the limit  $\mathbf{k}^2/a^2 \ll 1$  and let us disregard the gradient term. Under this condition we see that  $\phi_0$  and  $\delta\phi$  solve the same equation. The solutions have therefore to be related to each other by a constant of proportionality which depends upon time

$$\delta\phi = -\dot{\phi}_0 \delta t(\mathbf{x}). \quad (68)$$

This tells us that  $\phi(\mathbf{x}, t)$  will have the form

$$\phi(\mathbf{x}, t) = \phi_0(\mathbf{x}, t - \delta t(\mathbf{x})).$$

This equation indicates that the inflaton field does not acquire the same value at a given time  $t$  in all the space. On the contrary, when the inflaton field is rolling down its potential, it acquires different values from one spatial point  $\mathbf{x}$  to the next. The inflaton field is not homogeneous and fluctuations are present. These fluctuations, in turn, will induce fluctuations in the metric.

### 7.1 The metric fluctuations

The mathematical tool to describe the linear evolution of the cosmological perturbations is obtained by perturbing at the first order the FRW metric  $g_{\mu\nu}^{(0)}$ , see Eq. (1)

$$g_{\mu\nu} = g_{\mu\nu}^{(0)}(t) + g_{\mu\nu}(\mathbf{x}, t); \quad g_{\mu\nu} \ll g_{\mu\nu}^{(0)}. \quad (69)$$

The metric perturbations can be decomposed according to their spin with respect to a local rotation of the spatial coordinates on hypersurfaces of constant time. This leads to

- *scalar perturbations*
- *vector perturbations*
- *tensor perturbations*

Tensor perturbations or gravitational waves have spin 2 and are the true degrees of freedom of the gravitational fields in the sense that they can exist even in the vacuum. Vector perturbations are spin 1 modes arising from rotational velocity fields and are also called vorticity modes. Finally, scalar perturbations have spin 0.

Let us do a simple exercise to count how many scalar degrees of freedom are present. Take a space-time of dimensions  $D = n + 1$ , of which  $n$  coordinates are spatial coordinates. The symmetric metric tensor  $g_{\mu\nu}$  has  $\frac{1}{2}(n+2)(n+1)$  degrees of freedom. We can perform  $(n+1)$  coordinate transformations in order to eliminate  $(n+1)$  degrees of freedom, this leaves us with  $\frac{1}{2}n(n+1)$  degrees of freedom. These  $\frac{1}{2}n(n+1)$  degrees of freedom contain scalar, vector and tensor modes. According to Helmholtz's theorem we can always decompose a vector  $u_i$  ( $i = 1, \dots, n$ ) as  $u_i = \partial_i v + v_i$ , where  $v$  is a scalar (usually called potential flow) which is curl-free,  $v_{[i,j]} = 0$ , and  $v_i$  is a real vector (usually called vorticity) which is divergence-free,  $\nabla \cdot v = 0$ . This means that the real vector (vorticity) modes are  $(n-1)$ . Furthermore, a generic traceless tensor  $\Pi_{ij}$  can always be decomposed as  $\Pi_{ij} = \Pi_{ij}^S + \Pi_{ij}^V + \Pi_{ij}^T$ , where  $\Pi_{ij}^S = \left(-\frac{k_i k_j}{k^2} + \frac{1}{3}\delta_{ij}\right) \Pi$ ,  $\Pi_{ij}^V = (-i/2k)(k_i \Pi_j + k_j \Pi_i)$  ( $k_i \Pi_i = 0$ ) and  $k_i \Pi_{ij}^T = 0$ . This means that the true symmetric, traceless and transverse tensor degrees of freedom are  $\frac{1}{2}(n-2)(n+1)$ .

The number of scalar degrees of freedom is therefore

$$\frac{1}{2}n(n+1) - (n-1) - \frac{1}{2}(n-2)(n+1) = 2,$$

while the degrees of freedom of true vector modes are  $(n-1)$  and the number of degrees of freedom of true tensor modes (gravitational waves) is  $\frac{1}{2}(n-2)(n+1)$ . In four dimensions  $n = 3$ , meaning that one expects 2 scalar degrees of freedom, 2 vector degrees of freedom and 2 tensor degrees of freedom. As we shall see, to the 2 scalar degrees of freedom from the metric, one has to add another one, the inflaton field perturbation  $\delta\phi$ . However, since Einstein's equations will tell us that the two scalar degrees of freedom

from the metric are equal during inflation, we expect a total number of scalar degrees of freedom equal to 2.

At the linear order, the scalar, vector, and tensor perturbations evolve independently (they decouple) and it is therefore possible to analyse them separately. Vector perturbations are not excited during inflation because there are no rotational velocity fields during the inflationary stage. we shall analyse the generation of tensor modes (gravitational waves) in the following. For the time being we want to focus on the scalar degrees of freedom of the metric.

Considering only the scalar degrees of freedom of the perturbed metric, the most generic perturbed metric reads

$$g_{\mu\nu} = a^2 \begin{pmatrix} -1 - 2\Phi & \partial_i B \\ \partial_i B & (1 - 2\psi)\delta_{ij} + D_{ij}E \end{pmatrix}, \quad (70)$$

while the line-element can be written as

$$ds^2 = a^2((-1 - 2\Phi)d\tau^2 + 2\partial_i B d\tau dx^i + ((1 - 2\psi)\delta_{ij} + D_{ij}E) dx^i dx^j). \quad (71)$$

Here  $D_{ij} = (\partial_i \partial_j - \frac{1}{3} \delta_{ij} \nabla^2)$ .

## 7.2 The issue of gauge invariance

When studying the cosmological density perturbations, what we are interested in is following the evolution of a space-time which is neither homogeneous nor isotropic. This is done by following the evolution of the differences between the actual space-time and a well understood reference space-time. So we shall consider small perturbations away from the homogeneous, isotropic space-time.

The reference system in our case is the spatially flat Friedmann–Robertson–Walker (FRW) space-time, with line element  $ds^2 = a^2(\tau) \{d\tau^2 - \delta_{ij} dx^i dx^j\}$ . Now, the key issue is that general relativity is a gauge theory where the gauge transformations are the generic coordinate transformations from one local reference frame to another.

When we compute the perturbation of a given quantity, this is defined to be the difference between the value that this quantity assumes on the real physical space-time and the value it assumes on the unperturbed background. Nonetheless, to perform a comparison between these two values, it is necessary to compute them at the same space-time point. Since the two values live on two different geometries, it is necessary to specify a map which allows one to link univocally the same point on the two different space-times. This correspondence is called a gauge choice and changing the map means performing a gauge transformation.

Fixing a gauge in general relativity implies choosing a coordinate system. A choice of coordinates defines a *threading* of space-time into lines (corresponding to fixed spatial coordinates  $\mathbf{x}$ ) and a *slicing* into hypersurfaces (corresponding to fixed time  $\tau$ ). A choice of coordinates is called a *gauge* and there is no unique preferred gauge

GAUGE CHOICE  $\iff$  SLICING AND THREADING

From a more formal point of view, operating an infinitesimal gauge transformation on the coordinates

$$\widetilde{x}^\mu = x^\mu + \delta x^\mu \quad (72)$$

implies on a generic quantity  $Q$  a transformation on its perturbation

$$\delta \widetilde{Q} = \delta Q + \mathcal{L}_{\delta x} Q_0 \quad (73)$$



where  $Q_0$  is the value assumed by the quantity  $Q$  on the background and  $\mathcal{L}_{\delta x}$  is the Lie-derivative of  $Q$  along the vector  $\delta x^\mu$ .

Decomposing in the usual manner the vector  $\delta x^\mu$

$$\begin{aligned}\delta x^0 &= \xi^0(x^\mu); \\ \delta x^i &= \partial^i \beta(x^\mu) + v^i(x^\mu); \quad \partial_i v^i = 0,\end{aligned}\tag{74}$$

we can easily deduce the transformation law of a scalar quantity  $f$  (like the inflaton scalar field  $\phi$  and energy density  $\rho$ ). Instead of applying the formal definition (73), we find the transformation law in an alternative (and more pedagogical) way. We first write  $\delta f(x) = f(x) - f_0(x)$ , where  $f_0(x)$  is the background value. Under a gauge transformation we have  $\widetilde{\delta f}(\widetilde{x}^\mu) = \widetilde{f}(\widetilde{x}^\mu) - \widetilde{f}_0(\widetilde{x}^\mu)$ . Since  $f$  is a scalar we can write  $f(\widetilde{x}^\mu) = f(x^\mu)$  (the value of the scalar function in a given physical point is the same in all the coordinate system). On the other side, on the unperturbed background hypersurface  $\widetilde{f}_0 = f_0$ . We have therefore

$$\begin{aligned}\widetilde{\delta f}(\widetilde{x}^\mu) &= \widetilde{f}(\widetilde{x}^\mu) - \widetilde{f}_0(\widetilde{x}^\mu) \\ &= f(x^\mu) - f_0(\widetilde{x}^\mu) \\ &= f(\widetilde{x}^\mu) - f_0(\widetilde{x}^\mu) \\ &= f(\widetilde{x}^\mu) - \delta x^\mu \frac{\partial f}{\partial x^\mu}(\widetilde{x}) - f_0(\widetilde{x}^\mu),\end{aligned}\tag{75}$$

from which we finally deduce, being  $f_0 = f_0(x^0)$ ,

$$\widetilde{\delta f} = \delta f - f' \xi^0$$

For the spin-zero perturbations of the metric, we can proceed analogously. We use the following trick. Upon a coordinate transformation  $x^\mu \rightarrow \widetilde{x}^\mu = x^\mu + \delta x^\mu$ , the line element is left invariant,  $ds^2 = \widetilde{ds}^2$ . This implies, for instance, that  $a^2(\widetilde{x}^0) (1 + \widetilde{\Phi}) (\widetilde{dx}^0)^2 = a^2(x^0) (1 + \Phi) (dx^0)^2$ . Since  $a^2(\widetilde{x}^0) \simeq a^2(x^0) + 2a a' \xi^0$  and  $\widetilde{dx}^0 = (1 + \xi^{0'}) dx^0 + \frac{\partial x^0}{\partial x^i} dx^i$ , we obtain  $1 + 2\Phi = 1 + 2\widetilde{\Phi} + 2\mathcal{H}\xi^0 + 2\xi^{0'}$ . We now may introduce in detail some gauge-invariant quantities which play a major role in the computation of the density perturbations. In the following we shall be interested only in the coordinate transformations on constant time hypersurfaces and therefore gauge invariance will be equivalent to independence of the slicing.

### 7.3 The co-moving curvature perturbation

The intrinsic spatial curvature on hypersurfaces on constant conformal time  $\tau$  and for a flat universe is given by

$${}^{(3)}R = \frac{4}{a^2} \nabla^2 \psi.$$

The quantity  $\psi$  is usually referred to as the *curvature perturbation*. We have seen, however, that the curvature potential  $\psi$  is *not* gauge invariant, but is defined only on a given slicing. Under a transformation on constant time hypersurfaces  $t \rightarrow t + \delta\tau$  (change of the slicing)

$$\psi \rightarrow \psi + \mathcal{H} \delta\tau.$$

We now consider the *co-moving slicing* which is defined to be the slicing orthogonal to the worldlines of co-moving observers. The latter are free-falling and the expansion defined by them is isotropic. In practice, what this means is that there is no flux of energy measured by these observers, that is  $T_{0i} = 0$ . During inflation this means that these observers measure  $\delta\phi_{\text{com}} = 0$  since  $T_{0i}$  goes like  $\partial_i \delta\phi(\mathbf{x}, \tau) \phi'(\tau)$ .

Since  $\delta\phi \rightarrow \delta\phi - \phi' \delta\tau$  for a transformation on constant time hypersurfaces, this means that

$$\delta\phi \rightarrow \delta\phi_{\text{com}} = \delta\phi - \phi' \delta\tau = 0 \implies \delta\tau = \frac{\delta\phi}{\phi'},$$

that is  $\delta\tau = \frac{\delta\phi}{\phi'}$  is the time-displacement needed to go from a generic slicing with generic  $\delta\phi$  to the co-moving slicing where  $\delta\phi_{\text{com}} = 0$ . At the same time the curvature perturbation  $\psi$  transforms into

$$\psi \rightarrow \psi_{\text{com}} = \psi + \mathcal{H} \delta\tau = \psi + \mathcal{H} \frac{\delta\phi}{\phi'}.$$

The quantity

$$\mathcal{R} = \psi + \mathcal{H} \frac{\delta\phi}{\phi'} = \psi + H \frac{\delta\phi}{\dot{\phi}}$$

is the *co-moving curvature perturbation*. This quantity is gauge invariant by construction and is related to the gauge-dependent curvature perturbation  $\psi$  on a generic slicing to the inflaton perturbation  $\delta\phi$  in that gauge. By construction, the meaning of  $\mathcal{R}$  is that it represents the gravitational potential on co-moving hypersurfaces where  $\delta\phi = 0$  or the inflaton fluctuation hypersurfaces where  $\psi = 0$ :

$$\mathcal{R} = \psi|_{\delta\phi=0} = H \frac{\delta\phi}{\dot{\phi}} \Big|_{\psi=0}.$$

The power spectrum of the curvature perturbation may then be easily computed

$$\mathcal{R}_{\mathbf{k}} = H \frac{\delta\phi_{\mathbf{k}}}{\dot{\phi}}. \quad (76)$$

We may now compute the power spectrum of the co-moving curvature perturbation on superhorizon scales

$$\mathcal{P}_{\mathcal{R}}(k) = \frac{1}{2m_{\text{Pl}}^2 \epsilon} \left( \frac{H}{2\pi} \right)^2 \left( \frac{k}{aH} \right)^{n_{\mathcal{R}}-1} \equiv A_{\mathcal{R}}^2 \left( \frac{k}{aH} \right)^{n_{\mathcal{R}}-1}$$

where we have defined the *spectral index*  $n_{\mathcal{R}}$  of the co-moving curvature perturbation as

$$n_{\mathcal{R}} - 1 = \frac{d \ln \mathcal{P}_{\mathcal{R}}}{d \ln k} = 3 - 2\nu = 2\eta - 6\epsilon.$$

We conclude that inflation is responsible for the generation of adiabatic/curvature perturbations with an almost scale-independent spectrum. To compute the spectral index of the spectrum  $\mathcal{P}_{\mathcal{R}}(k)$  we have proceeded as follows: first solve the equation for the perturbation  $\delta\phi_{\mathbf{k}}$  in a de Sitter stage, with  $H = \text{constant}$  ( $\epsilon = \eta = 0$ ), whose solution is Eq. (56) and then taking into account the time-evolution of the Hubble rate and of  $\phi$  introducing the subscript in  $H_k$  and  $\phi_k$ . The time variation of the latter is determined by

$$\frac{d\ln \dot{\phi}_k}{d\ln k} = \left( \frac{d\ln \dot{\phi}_k}{dt} \right) \left( \frac{dt}{d\ln a} \right) \left( \frac{d\ln a}{d\ln k} \right) = \frac{\ddot{\phi}_k}{\dot{\phi}_k} \times \frac{1}{H} \times 1 = -\delta = \epsilon - \eta. \quad (77)$$

Correspondingly,  $\dot{\phi}_k$  is the value of the time derivative of the inflaton field when a given wavelength  $\sim k^{-1}$  crosses the horizon (from that point on the fluctuations remains frozen in). The curvature perturbation in such an approach would read

$$\mathcal{R}_{\mathbf{k}} \simeq \frac{H_k}{\dot{\phi}_k} \delta\phi_{\mathbf{k}} \simeq \frac{1}{2\pi} \left( \frac{H_k^2}{\dot{\phi}_k} \right).$$

Correspondingly

$$n_{\mathcal{R}} - 1 = \frac{d\ln \mathcal{P}_{\mathcal{R}}}{d\ln k} = \frac{d\ln H_k^4}{d\ln k} - \frac{d\ln \dot{\phi}_k^2}{d\ln k} = -4\epsilon + (2\eta - 2\epsilon) = 2\eta - 6\epsilon.$$

During inflation the curvature perturbation is generated on superhorizon scales with a spectrum which is nearly scale invariant [13], that is, is nearly independent of the wavelength  $\lambda = \pi/k$ : the amplitude of the fluctuation on superhorizon scales does not (almost) depend upon the time at which the fluctuation crosses the horizon and becomes frozen in. The small tilt of the power spectrum arises from the fact that the inflaton field is massive, giving rise to a non-vanishing  $\eta$  and because during inflation the Hubble rate is not exactly constant, but nearly constant, where ‘nearly’ is quantified by the slow-roll parameters  $\epsilon$ .

*Comment:* From what we have found so far, we may conclude that on superhorizon scales the co-moving curvature perturbation  $\mathcal{R}$  and the uniform-density gauge curvature  $\zeta$  satisfy on superhorizon scales the relation

$$\dot{\mathcal{R}}_{\mathbf{k}} \simeq 0.$$

## 7.4 Gravitational waves

Quantum fluctuations in the gravitational fields are generated in a similar fashion to that of the scalar perturbations discussed so far. A gravitational wave may be viewed as a ripple of space-time in the FRW background metric (1) and in general the linear tensor perturbations may be written as

$$g_{\mu\nu} = a^2(\tau) \left[ -d\tau^2 + (\delta_{ij} + h_{ij}) dx^i dx^j \right],$$

where  $|h_{ij}| \ll 1$ . The tensor  $h_{ij}$  has six degrees of freedom, but, as we studied in Subsection 7.1, the tensor perturbations are traceless,  $\delta^{ij} h_{ij} = 0$ , and transverse  $\partial^i h_{ij} = 0$  ( $i = 1, 2, 3$ ). With these four constraints, there remain two physical degrees of freedom, or polarizations, which are usually indicated  $\lambda = +, \times$ . More precisely, we can write

$$h_{ij} = h_+ e_{ij}^+ + h_{\times} e_{ij}^{\times},$$

where  $e^+$  and  $e^{\times}$  are the polarization tensors which have the following properties

$$e_{ij} = e_{ji}, \quad k^i e_{ij} = 0, \quad , e_{ii} = 0,$$

$$e_{ij}(-\mathbf{k}, \lambda) = e_{ij}^*(\mathbf{k}, \lambda), \quad \sum_{\lambda} e_{ij}^*(\mathbf{k}, \lambda) e^{ij}(\mathbf{k}, \lambda) = 4.$$

Notice also that the tensors  $h_{ij}$  are gauge-invariant and therefore represent physical degrees of freedom.

If the stress-energy momentum tensor is diagonal, as the one provided by the inflaton potential  $T_{\mu\nu} = \partial_{\mu}\phi\partial_{\nu}\phi - g_{\mu\nu}\mathcal{L}$ , the tensor modes do not have any source in their equation and their action can be written as

$$\frac{m_{\text{Pl}}^2}{2} \int d^4x \sqrt{-g} \frac{1}{2} \partial_{\sigma} h_{ij} \partial^{\sigma} h_{ij},$$

that is the action of four independent massless scalar fields. The gauge-invariant tensor amplitude

$$v_{\mathbf{k}} = am_{\text{Pl}} \frac{1}{\sqrt{2}} h_{\mathbf{k}},$$

satisfies therefore the equation

$$v_{\mathbf{k}}'' + \left( k^2 - \frac{a''}{a} \right) v_{\mathbf{k}} = 0,$$

which is the equation of motion of a massless scalar field in a quasi-de Sitter epoch. We can therefore make use of the results present in Subsection 6.5 and Eq. (63) to conclude that on superhorizon scales the tensor modes scale like

$$|v_{\mathbf{k}}| = \left( \frac{H}{2\pi} \right) \left( \frac{k}{aH} \right)^{\frac{3}{2} - \nu_T},$$

where

$$\nu_T \simeq \frac{3}{2} - \epsilon.$$

Since fluctuations are (nearly) frozen in on superhorizon scales, a way of characterizing the tensor perturbations is to compute the spectrum on scales larger than the horizon

$$\mathcal{P}_T(k) = \frac{k^3}{2\pi^2} \sum_{\lambda} |h_{\mathbf{k}}|^2 = 4 \times 2 \frac{k^3}{2\pi^2} |v_{\mathbf{k}}|^2. \quad (78)$$

This gives the power spectrum on superhorizon scales

$$\mathcal{P}_T(k) = \frac{8}{m_{\text{Pl}}^2} \left( \frac{H}{2\pi} \right)^2 \left( \frac{k}{aH} \right)^{n_T} \equiv A_T^2 \left( \frac{k}{aH} \right)^{n_T}$$

where we have defined the *spectral index*  $n_T$  of the tensor perturbations as

$$n_T = \frac{d \ln \mathcal{P}_T}{d \ln k} = 3 - 2\nu_T = -2\epsilon.$$

The tensor perturbation is almost scale-invariant. Notice that the amplitude of the tensor modes depends only on the value of the Hubble rate during inflation. This amounts to saying that it depends only on the energy scale  $V^{1/4}$  associated to the inflaton potential. A detection of gravitational waves from inflation will therefore be a direct measurement of the energy scale associated to inflation.

### 7.5 The consistency relation

The results obtained so far for the scalar and tensor perturbations allow one to predict a *consistency relation* which holds for the models of inflation addressed in these lectures, i.e., the models of inflation driven by one-single field  $\phi$ . We define the tensor-to-scalar amplitude ratio to be

$$r = \frac{\frac{1}{100} A_T^2}{\frac{4}{25} A_{\mathcal{R}}^2} = \frac{\frac{1}{100} 8 \left( \frac{H}{2\pi m_{\text{Pl}}} \right)^2}{\frac{4}{25} (2\epsilon)^{-1} \left( \frac{H}{2\pi m_{\text{Pl}}} \right)^2} = \epsilon.$$

This means that

$$r = -\frac{n_T}{2}$$

One-single models of inflation predict that during inflation driven by a single scalar field, the ratio between the amplitude of the tensor modes and that of the curvature perturbations is equal to minus one-half of the tilt of the spectrum of tensor modes. If this relation turns out to be falsified by the future measurements of the CMB anisotropies, this does not mean that inflation is wrong, but only that inflation has not been driven by only one field.

### 7.6 From the inflationary seeds to the matter power spectrum

As the curvature perturbations enter the causal horizon during radiation- or matter-domination, they create density fluctuations  $\delta\rho_{\mathbf{k}}$  via gravitational attractions of the potential wells. The density contrast  $\delta_{\mathbf{k}} = \frac{\delta\rho_{\mathbf{k}}}{\bar{\rho}}$  can be deduced from the Poisson equation

$$\frac{k^2 \Phi_{\mathbf{k}}}{a^2} = -4\pi G \delta\rho_{\mathbf{k}} = -4\pi G \frac{\delta\rho_{\mathbf{k}}}{\bar{\rho}} \bar{\rho} = \frac{3}{2} H^2 \frac{\delta\rho_{\mathbf{k}}}{\bar{\rho}}$$

where  $\bar{\rho}$  is the background average energy density. This means that

$$\delta_{\mathbf{k}} = \frac{2}{3} \left( \frac{k}{aH} \right)^2 \Phi_{\mathbf{k}}.$$

From this expression we can compute the power spectrum of matter density perturbations induced by inflation when they re-enter the horizon during matter-domination:

$$\mathcal{P}_{\delta\rho} = \langle |\delta_{\mathbf{k}}|^2 \rangle = A \left( \frac{k}{aH} \right)^n = \frac{2\pi^2}{k^3} \left( \frac{2}{5} \right)^2 A_{\mathcal{R}}^2 \left( \frac{k}{aH} \right)^4 \left( \frac{k}{aH} \right)^{n_{\mathcal{R}}-1}$$

from which we deduce that matter perturbations scale linearly with the wave-number and have a scalar tilt

$$n = n_{\mathcal{R}} = 1 + 2\eta - 6\epsilon.$$

The primordial spectrum  $\mathcal{P}_{\delta\rho}$  is of course reprocessed by gravitational instabilities after the universe becomes matter-dominated. Indeed, as we have seen in Section 6, perturbations evolve after entering the horizon and the power spectrum will not remain constant. To see how the density contrast is reprocessed we have first to analyse how it evolves on superhorizon scales before horizon-crossing. We use the following trick. Consider a flat universe with average energy density  $\bar{\rho}$ . The corresponding Hubble rate is

$$H^2 = \frac{8\pi G}{3} \bar{\rho}.$$

A small positive fluctuation  $\delta\rho$  will cause the universe to be closed:

$$H^2 = \frac{8\pi G}{3} (\bar{\rho} + \delta\rho) - \frac{k}{a^2}.$$

Subtracting the two equations we find

$$\frac{\delta\rho}{\rho} = \frac{3}{8\pi G} \frac{k}{a^2 \rho} \sim \begin{cases} a^2 & \text{RD} \\ a & \text{MD} \end{cases}$$

Notice that  $\Phi_{\mathbf{k}} \sim \delta\rho a^2/k^2 \sim (\delta\rho/\rho) \rho a^2/k^2 = \text{constant}$  for both RD and MD which confirms our previous findings.

When the matter densities enter the horizon, they do not increase appreciably before matter-domination because before matter-domination pressure is too large and does not allow the matter inhomogeneities to grow. On the other hand, the suppression of growth due to radiation is restricted to scales smaller than the horizon, while large-scale perturbations remain unaffected. This is why the horizon size at equality sets an important scale for structure growth:

$$k_{\text{EQ}} = H^{-1}(a_{\text{EQ}}) \simeq 0.08 h \text{ Mpc}^{-1}.$$

Therefore, perturbations with  $k \gg k_{\text{EQ}}$  are perturbations which have entered the horizon before matter-domination and have remained nearly constant till equality. This means that they are suppressed with respect to those perturbations having  $k \ll k_{\text{EQ}}$  by a factor  $(a_{\text{ENT}}/a_{\text{EQ}})^2 = (k_{\text{EQ}}/k)^2$ . If we define the transfer function  $T(k)$  by the relation  $\mathcal{R}_{\text{final}} = T(k) \mathcal{R}_{\text{initial}}$  we find therefore that roughly speaking

$$T(k) = \begin{cases} 1 & k \ll k_{\text{EQ}}, \\ (k_{\text{EQ}}/k)^2 & k \gg k_{\text{EQ}}. \end{cases}$$

The corresponding power spectrum will be

$$\mathcal{P}_{\delta\rho}(k) \sim \begin{cases} \left(\frac{k}{aH}\right) & k \ll k_{\text{EQ}}, \\ \left(\frac{k}{aH}\right)^{-3} & k \gg k_{\text{EQ}}. \end{cases}$$

Of course, a more careful computation needs to include many other effects such as neutrino free-streaming, photon diffusion and the diffusion of baryons along with photons. It is encouraging, however, that this rough estimate turns out to be confirmed by present data on large-scale structures [4].

The next step would be to investigate how the primordial perturbations generated by inflation flow into the CMB to produce their anisotropies.

## 8 From inflation to large-angle CMB anisotropy

As we have already mentioned, the high temperature of the early universe maintained a low equilibrium fraction of neutral atoms, and a correspondingly high number density of free electrons. Coulomb scattering between the ions and electrons kept them in local kinetic equilibrium, and Thomson scattering of

photons tended to maintain the isotropy of the CMB in the baryon rest frame. As the universe expanded and cooled, the dominant element hydrogen started to recombine when the temperature fell below  $\sim 4000$  K. This is a factor of 40 lower than might be anticipated from the 13.6 eV ionization potential of hydrogen, and is due to the large ratio of the number of photons to baryons. Through recombination, the mean-free path for Thomson scattering grew to the horizon size and CMB photons “decoupled” from matter. More precisely, the probability density that photons last scattered at some time defines the visibility function. This is now known to peak 380 kyr after the Big Bang with a width  $\sim 120$  kyr. Since then, CMB photons have propagated relatively unimpeded for 13.7 Gyr, covering a co-moving distance  $\sim 14.1$  Gpc. The distribution of their energies carries the imprint of fluctuations in the radiation temperature, the gravitational potentials, and the peculiar velocity of the radiation where they last scattered, as the temperature anisotropies that we observe today.

Temperature fluctuations in the CMB arise due to various distinct physical effects: our peculiar velocity with respect to the cosmic rest frame; fluctuations in the gravitational potential on the last scattering surface; fluctuations intrinsic to the radiation field itself on the last scattering surface; the peculiar velocity of the last scattering surface and damping of anisotropies if the universe should be re-ionized after decoupling. The first effect gives rise to the dipole anisotropy. Finally, there is the contribution from the evolution of the anisotropies from the last scattering surface till today (which we shall neglect from now on).

The second effect, known as the Sachs–Wolfe effect is the dominant contribution to the anisotropy on large-angular scales,  $\theta \gg \theta_{\text{HOR}} \sim 1^\circ$ . The last three effects provide the dominant contributions to the anisotropy on small-angular scales,  $\theta \ll 1^\circ$ .

## 8.1 Sachs–Wolfe plateau

We consider first the temperature fluctuations on large-angular scales that arise due to the Sachs–Wolfe effect. These anisotropies probe length scales that were superhorizon-sized at photon decoupling and therefore insensitive to microphysical processes. On the contrary, they provide a probe of the original spectrum of primeval fluctuations produced during inflation.

To proceed, we consider the CMB anisotropy measured at positions other than our own and at earlier times. This is called the brightness function  $\Theta(t, \mathbf{x}, \mathbf{n}) \equiv \delta T(t, \mathbf{x}, \mathbf{n})/T(t)$ . The photons with momentum  $\mathbf{p}$  in a given range  $d^3p$  have intensity  $I$  proportional to  $T^4(t, \mathbf{x}, \mathbf{n})$  and therefore  $\delta I/I = 4\Theta$ . The brightness function depends upon the direction  $\mathbf{n}$  of the photon momentum or, equivalently, on the direction of observation  $\mathbf{e} = -\mathbf{n}$ . Because the CMB travels freely from the last-scattering, we can write

$$\frac{\delta T}{T} = \Theta(t_{\text{LS}}, \mathbf{x}_{\text{LS}}, \mathbf{n}) + \left( \frac{\delta T}{T} \right)_*,$$

where  $\mathbf{x}_{\text{LS}} = -x_{\text{LS}}\mathbf{n}$  is the point of the origin of the photon coming from the direction  $\mathbf{e}$ . The co-moving distance of the last scattering distance is  $x_{\text{LS}} = 2/H_0$ . The first term corresponds to the anisotropy already present at last scattering and the second term is the additional anisotropy acquired during the travel towards us, equal to minus the fractional perturbation in the redshift of the radiation. Notice that the separation between each term depends on the slicing, but the sum does not.

Consider the redshift perturbation on co-moving slicing. We imagine the universe populated by co-moving observers along the line of sight. The relative velocity of adjacent co-moving observers is equal to their distance times the velocity gradient measured along  $\mathbf{n}$  of the photon. In the unperturbed universe, we have  $\mathbf{u} = H\mathbf{r}$ , leading to the velocity gradient  $u_{ij} = \partial u_i / \partial r_j = u_{ij} = H(t)\delta_{ij}$  with zero vorticity and shear. Including a peculiar velocity field as perturbation,  $\mathbf{u} = H\mathbf{r} + \mathbf{v}$  and  $u_{ij} = H(t)\delta_{ij} + \frac{1}{a} \frac{\partial v_i}{\partial x_j}$ . The corresponding Doppler shift is

$$\frac{d\lambda}{\lambda} = \frac{da}{a} + n_i n_j \frac{\partial v_i}{\partial x_j} dx.$$

The perturbed FRW equation is

$$\delta H = \frac{1}{3} \nabla \cdot \mathbf{v},$$

while

$$(\delta \rho)' = -3\rho\delta H - 3H\delta\rho.$$

Instead of  $\delta\rho$ , let us work with the density contrast  $\delta = \delta\rho/\rho$ . Remembering that  $\rho \sim a^{-3}$ , we find that  $\dot{\delta} = -3\delta H$ , which gives

$$\nabla \cdot \mathbf{v} = -\dot{\delta}_{\mathbf{k}}.$$

From the Euler equation  $\dot{\mathbf{u}} = -\rho^{-1}\nabla p - \nabla\Phi$ , we deduce  $\dot{\mathbf{v}} + H\mathbf{v} = -\nabla\Phi - \rho^{-1}\nabla p$ . Therefore, for  $a \sim t^{2/3}$  and negligible pressure gradient, since the gravitational potential is constant, we find

$$\mathbf{v} = -t\nabla\Phi$$

leading to

$$\left(\frac{\delta T}{T}\right)_* = \int_0^{x_{\text{LS}}} \frac{t}{a} \frac{d^2\Phi}{dx^2} dx. \quad (79)$$

The photon trajectory is  $ad\mathbf{x}/dt = \mathbf{n}$ . Using  $a \sim t^{2/3}$  gives

$$x(t) = \int_t^{t_0} \frac{dt'}{a} = 3 \left( \frac{a_0}{t_0} - \frac{t}{a} \right).$$

Integrating by parts Eq. (79), we finally find

$$\left(\frac{\delta T}{T}\right)_* = \frac{1}{3} [\Phi(\mathbf{x}_{\text{LS}}) - \Phi(0)] + \mathbf{e} \cdot [\mathbf{v}(0, t_0) - \mathbf{v}(\mathbf{x}_{\text{LS}}, t_{\text{LS}})].$$

The potential at our position contributes only to the unobservable monopole and can be dropped. On scales outside the horizon,  $\mathbf{v} = -t\nabla\Phi \sim 0$ . The remaining term is the Sachs–Wolfe effect

$$\frac{\delta T(\mathbf{e})}{T} = \frac{1}{3} \Phi(\mathbf{x}_{\text{LS}}) = \frac{1}{5} \mathcal{R}(\mathbf{x}_{\text{LS}}).$$

This relation has been obtained as follows. The co-moving curvature perturbation is given during the radiation phase by  $\mathcal{R} = \psi + H\delta\rho/\dot{\rho} = \psi - 1/3\delta\rho_\gamma/\rho_\gamma$ . Einstein equations set  $\psi = \Phi$  and  $\delta\rho_\gamma/\rho_\gamma = -2\Phi$  on super-horizon scales. Therefore  $\mathcal{R} = 5/3\Phi$  beyond the horizon.

At large angular scales, the theory of cosmological perturbations predicts a remarkably simple formula relating the CMB anisotropy to the curvature perturbation generated during inflation.

We have seen previously that the temperature anisotropy is commonly expanded in spherical harmonics  $\frac{\Delta T}{T}(x_0, \tau_0, \mathbf{n}) = \sum_{\ell m} a_{\ell, m}(x_0) Y_{\ell m}(\mathbf{n})$ , where  $x_0$  and  $\tau_0$  are our position and the preset time, respectively,  $\mathbf{n}$  is the direction of observation,  $\ell$ 's are the different multipoles, and  $\langle a_{\ell m} a_{\ell' m'}^* \rangle = \delta_{\ell, \ell'} \delta_{m, m'} C_\ell$ , where the deltas are due to the fact that the process that created the anisotropy is statistically isotropic. The  $C_\ell$ 's are the so-called CMB power spectrum. For homogeneity and isotropy, the  $C_\ell$ 's are neither a function of  $x_0$ , nor of  $m$ . The two-point-correlation function is related to the  $C_\ell$ 's according to Eq. (23).



For adiabatic perturbations we have seen that on large scales, larger than the horizon on the last scattering surface (corresponding to angles larger than  $\theta_{\text{HOR}} \sim 1^\circ$ )  $\delta T/T = \frac{1}{3}\Phi(\mathbf{x}_{\text{LS}})$ . In Fourier transform

$$\frac{\delta T(\mathbf{k}, \tau_0, \mathbf{n})}{T} = \frac{1}{3}\Phi_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{n}(\tau_0 - \tau_{\text{LS}})}. \quad (80)$$

Using the decomposition

$$\exp(i\mathbf{k}\cdot\mathbf{n}(\tau_0 - \tau_{\text{LS}})) = \sum_{\ell=0}^{\infty} (2\ell+1) i^\ell j_\ell(k(\tau_0 - \tau_{\text{LS}})) P_\ell(\mathbf{k}\cdot\mathbf{n}) \quad (81)$$

where  $j_\ell$  is the spherical Bessel function of order  $\ell$  and substituting, we get

$$\begin{aligned} \left\langle \frac{\delta T(x_0, \tau_0, \mathbf{n})}{T} \frac{\delta T(x_0, \tau_0, \mathbf{n}')}{T} \right\rangle &= \\ &= \frac{1}{V} \int d^3x \left\langle \frac{\delta T(x_0, \tau_0, \mathbf{n})}{T} \frac{\delta T(x_0, \tau_0, \mathbf{n}')}{T} \right\rangle = \\ &= \frac{1}{(2\pi)^3} \int d^3k \left\langle \frac{\delta T(\mathbf{k}, \tau_0, \mathbf{n})}{T} \left( \frac{\delta T(\mathbf{k}, \tau_0, \mathbf{n}')}{T} \right)^* \right\rangle = \\ &= \frac{1}{(2\pi)^3} \int d^3k \left( \left\langle \frac{1}{3} |\Phi|^2 \right\rangle \sum_{\ell, \ell'=0}^{\infty} (2\ell+1)(2\ell'+1) j_\ell(k(\tau_0 - \tau_{\text{LS}})) \right. \\ &\quad \left. j_{\ell'}(k(\tau_0 - \tau_{\text{LS}})) P_\ell(\mathbf{k}\cdot\mathbf{n}) P_{\ell'}(\mathbf{k}'\cdot\mathbf{n}') \right) \end{aligned} \quad (82)$$

Inserting  $P_\ell(\mathbf{k}\cdot\mathbf{n}) = \frac{4\pi}{2\ell+1} \sum_m Y_{lm}^*(\mathbf{k}) Y_{\ell m}(\mathbf{n})$  and analogously for  $P_\ell(\mathbf{k}'\cdot\mathbf{n}')$ , integrating over the directions  $d\Omega_k$  generates  $\delta_{\ell\ell'} \delta_{mm'} \sum_m Y_{\ell m}^*(\mathbf{n}) Y_{\ell m}(\mathbf{n}')$ . Using as well  $\sum_m Y_{\ell m}^*(\mathbf{n}) Y_{\ell m}(\mathbf{n}') = \frac{2\ell+1}{4\pi} P_\ell(\mathbf{n}\cdot\mathbf{n}')$ , we get

$$\begin{aligned} &\left\langle \frac{\delta T(x_0, \tau_0, \mathbf{n})}{T} \frac{\delta T(x_0, \tau_0, \mathbf{n}')}{T} \right\rangle \\ &= \sum_\ell \frac{2\ell+1}{4\pi} P_\ell(\mathbf{n}\cdot\mathbf{n}') \frac{2}{\pi} \int \frac{dk}{k} \left\langle \frac{1}{9} |\Phi|^2 \right\rangle k^3 j_\ell^2(k(\tau_0 - \tau_{\text{LS}})). \end{aligned} \quad (84)$$

Comparing this expression with that for the  $C_\ell$ , we get the expression for the  $C_\ell^{\text{AD}}$ , where the suffix “AD” stands for adiabatic:

$$C_\ell^{\text{AD}} = \frac{2}{\pi} \int \frac{dk}{k} \left\langle \frac{1}{9} |\Phi|^2 \right\rangle k^3 j_\ell^2(k(\tau_0 - \tau_{\text{LS}})) \quad (85)$$

which is valid for  $2 \leq \ell \ll (\tau_0 - \tau_{\text{LS}})/\tau_{\text{LS}} \sim 100$ .

If we generically indicate by  $\langle |\Phi_{\mathbf{k}}|^2 \rangle k^3 = A^2 (k\tau_0)^{n-1}$ , we can perform the integration and get

$$\frac{\ell(\ell+1)C_\ell^{\text{AD}}}{2\pi} = \left[ \frac{\sqrt{\pi}}{2} \ell(\ell+1) \frac{\Gamma(\frac{3-n}{2}) \Gamma(\ell + \frac{n-1}{2})}{\Gamma(\frac{4-n}{2}) \Gamma(\ell + \frac{5-n}{2})} \right] \frac{A^2}{9} \left( \frac{H_0}{2} \right)^{n-1}. \quad (86)$$

For  $n \simeq 1$  and  $\ell \gg 1$ , we can approximate this expression to

$$\frac{\ell(\ell+1)C_\ell^{\text{AD}}}{2\pi} = \frac{A^2}{9}. \quad (87)$$

This result shows that inflation predicts a very flat spectrum for low  $\ell$ . Furthermore, since inflation predicts  $\Phi_{\mathbf{k}} = \frac{3}{5}\mathcal{R}_{\mathbf{k}}$ , we find that

$$\pi \ell(\ell+1)C_\ell^{\text{AD}} = \frac{A_{\mathcal{R}}^2}{25} = \frac{1}{25} \frac{1}{2 m_{\text{Pl}}^2 \epsilon} \left( \frac{H}{2\pi} \right)^2. \quad (88)$$

WMAP5 data imply that  $\frac{\ell(\ell+1)C_\ell^{\text{AD}}}{2\pi} \simeq 10^{-10}$  or

$$\left(\frac{V}{\epsilon}\right)^{1/4} \simeq 6.7 \times 10^{16} \text{ GeV}$$

## 8.2 Acoustic peaks

To be able to calculate the power spectrum of the anisotropies even on angular scales larger than  $1^\circ$ , we need to consider the evolution of the photon anisotropies. As we already mentioned, before recombination Thomson scattering was very efficient. As a result it is a good approximation to treat photons and baryons as a single fluid. This treatment is called the tight-coupling approximation and will allow us to evolve the perturbations until recombination.

The equation for the photon density perturbations for one Fourier mode of wave-number  $k$  is that of a forced and damped harmonic oscillator

$$\begin{aligned} \ddot{\delta}_\gamma + \frac{\dot{R}}{(1+R)}\dot{\delta}_\gamma + k^2 c_s^2 \delta_\gamma &= F, \\ F &= 4[\ddot{\psi} + \frac{\dot{R}}{(1+R)}\dot{\psi} - \frac{1}{3}k^2\Phi], \\ \dot{\delta}_\gamma &= -\frac{4}{3}kv_\gamma + 4\dot{\psi}. \end{aligned} \tag{89}$$

The photon–baryon fluid can sustain acoustic oscillations. The inertia is provided by the baryons, while the pressure is provided by the photons. The sound speed is  $c_s^2 = 1/3(1+R)$ , with  $R = 3\rho_b/4\rho_\gamma = 31.5 (\Omega_b h^2)(T/2.7)^{-4}[(1+z)/10^3]^{-1}$ . As the baryon fraction goes down, the sound speed approaches  $c_s^2 \rightarrow 1/3$ . The third equation above is the continuity equation.

As a toy problem, we shall solve Eq. (89) under some simplifying assumptions. If we consider a matter-dominated universe, the driving force becomes a constant,  $F = -4/3k^2\Phi$ , because the gravitational potential remains constant in time. We neglect anisotropic stresses so that  $\psi = \Phi$ , and, furthermore, we neglect the time dependence of  $R$ . Equation (89) becomes that of a harmonic oscillator that can be trivially solved. This is a very simplified picture, but it captures most of the relevant physics we want to discuss.

To obtain the final solution we need again to specify the initial conditions. we shall restrict ourselves to adiabatic initial conditions, the most natural outcome of inflation. In our context this means that initially  $\Phi = \psi = \Phi_0$ ,  $\delta_\gamma = -8/3\Phi_0$ , and  $v_\gamma = 0$ . We have denoted  $\Phi_0$  the initial amplitude of the potential fluctuations. We shall take  $\Phi_0$  to be a Gaussian random variable with power spectrum  $P_{\Phi_0}$ .

We have made enough approximations that the evaluation of the sources in the integral solution has become trivial. The solution for the density and velocity of the photon fluid at recombination is

$$\begin{aligned} \left(\frac{\delta_\gamma}{4} + \Phi\right)|_{\text{LS}} &= \frac{\Phi_0}{3}(1+3R)\cos(kc_s\tau_{\text{LS}}) - \Phi_0 R, \\ v_\gamma|_{\tau_{\text{LS}}} &= -\Phi_0(1+3R)c_s\sin(kc_s\tau_{\text{LS}}). \end{aligned} \tag{90}$$

Equation (90) is the solution for a single Fourier mode. All quantities have an additional spatial dependence ( $e^{i\mathbf{k}\cdot\mathbf{x}}$ ), which we have not included in order to make the notation more compact. With that additional term the solution we have is

$$\begin{aligned} \frac{\delta T}{T}(\mathbf{n}) &= e^{ikD_{\text{LS}}\cos\theta} S \\ S &= \Phi_0 \frac{(1+3R)}{3} [\cos(kc_s\tau_{\text{LS}}) - \frac{3R}{(1+3R)}], \end{aligned}$$

$$-i\sqrt{\frac{3}{1+R}}\cos\theta\sin(kc_s\tau_{\text{LS}})], \quad (91)$$

where we have neglected the  $\Phi$  on the left-hand side because it is a constant. We have introduced  $\cos\theta$ , the cosine of the angle between the direction of observation and the wavevector  $\mathbf{k}$ ; for example,  $\mathbf{k} \cdot \mathbf{x} = kD_{\text{LS}}\cos\theta$ . The term proportional to  $\cos\theta$  is the Doppler contribution.

Once the temperature perturbation produced by one Fourier mode has been calculated, we need to expand it into spherical harmonics. The power spectrum of temperature anisotropies is expressed in terms of the  $a_{\ell m}$  coefficients as  $C_{T\ell} = \sum_m |a_{\ell m}|^2$ . The contribution to  $C_{T\ell}$  from each Fourier mode is weighted by the amplitude of primordial fluctuations in this mode, characterized by the power spectrum of  $\Phi_0 = 3/5\mathcal{R}$ ,  $P_{\Phi_0} = Ak^{-3}$  as dictated by inflation. In practice, fluctuations on angular scale  $\ell$  receive most of their contributions from wavevectors around  $k_* = \ell/D_{\text{LS}}$ , so roughly the amplitude of the power spectrum at multipole  $\ell$  is given by the value of the sources in Eq. (90) at  $k_*$ .

After summing the contributions from all modes, the power spectrum is roughly given by

$$\begin{aligned} \ell(\ell+1)C_{T\ell} \approx & A\left\{\left[\frac{(1+3R)}{3}\cos(k_*c_s\tau_{\text{LS}}) - R\right]^2 + \right. \\ & \left. \frac{(1+3R)^2}{3}c_s^2\sin^2(k_*c_s\tau_{\text{LS}})\right\}. \end{aligned} \quad (92)$$

Equation (92) can be used to understand the basic features in the CMB power spectra. The baryon drag on the photon–baryon fluid reduces its sound speed below  $1/3$  and makes the monopole contribution dominant (the one proportional to  $\cos(k_*c_s\tau_{\text{LS}})$ ). Thus, the  $C_{T\ell}$  spectrum peaks where the monopole term peaks,  $k_*c_s\tau_{\text{LS}} = \pi, 2\pi, 3\pi, \dots$ , which correspond to  $\ell_{\text{peak}} = n\pi D_{\text{LS}}/c_s\tau_{\text{LS}}$ .

It is very important to understand the origin of the acoustic peaks. In this model the universe is filled with standing waves; all modes of wave-number  $k$  are in phase, which leads to the oscillatory terms. The sine and cosine in Eq. (92) originate in the time dependence of the modes. Each mode  $\ell$  receives contributions preferentially from Fourier modes of a particular wavelength  $k_*$  (but pointing in all directions), so to obtain peaks in  $C_\ell$ , it is crucial that all modes of a given  $k$  be in phase. If this is not the case, the features in the  $C_{T\ell}$  spectra will be blurred and can even disappear. This is what happens when one considers the spectra produced by topological defects. The phase coherence of all modes of a given wave-number can be traced to the fact that perturbations were produced very early on and had wavelengths larger than the horizon during many expansion times.

There are additional physical effects we have neglected. The universe was radiation dominated early on, and modes of wavelength smaller and bigger than the horizon at matter–radiation equality behave differently. During the radiation era the perturbations in the photon–baryon fluid are the main source for the gravitational potentials which decay once a mode enters into the horizon. The gravitational potential decay acts as a driving force for the oscillator in Eq. (89), so a feedback loop is established. As a result, the acoustic oscillations for modes that entered the horizon before matter–radiation equality have a higher amplitude. In the  $C_{T\ell}$  spectrum the separation between modes that experience this feedback and those that do not occurs at  $\ell \sim D_{\text{LS}}/\tau_{\text{LS}}$ . Larger  $\ell$  values receive their contributions from modes that entered the horizon before matter–radiation equality. Finally, when a mode is inside the horizon during the radiation era the gravitational potentials decay.

There is a competing effect, Silk damping, that reduces the amplitude of the large- $\ell$  modes. The photon–baryon fluid is not a perfect fluid. Photons have a finite mean free path and thus can random-walk away from the peaks and valleys of the standing waves. Thus perturbations of wavelength comparable to or smaller than the distance the photons can random-walk get damped. This effect can be modelled by multiplying Eq. 91 by  $\exp(-k^2/k_s^2)$ , with  $k_s^{-1} \propto \tau_{\text{LS}}^{1/2}(\Omega_b h^2)^{-1/2}$ . Silk damping is important for multipoles of order  $\ell_{\text{Silk}} \sim k_s D_{\text{LS}}$ . Finally, the last scattering surface has a finite width. Perturbations

with wavelength comparable to this width get smeared out due to cancellations along the line of sight. This effect introduces an additional damping with a characteristic scale  $k_w^{-1} \propto \delta\tau_{LS}$ .

The location of the first peak is by itself a measurement of the geometry of the universe. In fact, photons propagating on geodesics from the last scattering surface to us feel the spatial geometry, whose properties we learned are dictated by  $\Omega_0$ . In fact, the location of the first peak is given by  $\ell_1 \simeq 220/\sqrt{\Omega_0}$ . WMAP5 gives  $\Omega_0 = 1.00^{+0.07}_{-0.03}$ . This tells us that the spatial (local) geometry of the universe is flat. This is precisely what inflation predicts.

### 8.3 The polarization of the CMB anisotropies

The anisotropy field is characterized by a  $2 \times 2$  intensity tensor  $I_{ij}$ . For convenience, we normalize this tensor so that it represents the fluctuations in units of the mean intensity ( $I_{ij} = \delta I/I_0$ ). The intensity tensor is a function of direction on the sky,  $\mathbf{n}$ , and two directions perpendicular to  $\mathbf{n}$  that are used to define its components ( $\mathbf{e}_1, \mathbf{e}_2$ ). The Stokes parameters  $Q$  and  $U$  are defined as  $Q = (I_{11} - I_{22})/4$  and  $U = I_{12}/2$ , while the temperature anisotropy is given by  $T = (I_{11} + I_{22})/4$  (the factor of 4 relates fluctuations in the intensity with those in the temperature,  $I \propto T^4$ ). When representing polarization using “rods” in a map, the magnitude is given by  $P = \sqrt{Q^2 + U^2}$ , and the orientation makes an angle  $\alpha = \frac{1}{2} \arctan(U/Q)$  with  $\mathbf{e}_1$ . In principle the fourth Stokes parameter  $V$  that describes circular polarization is needed, but we ignore it because it cannot be generated through Thomson scattering, so the CMB is not expected to be circularly polarized. While the temperature is invariant under a right-handed rotation in the plane perpendicular to direction  $\mathbf{n}$ ,  $Q$  and  $U$  transform under rotation by an angle  $\psi$  as

$$(Q \pm iU)'(\mathbf{n}) = e^{\mp 2i\psi}(Q \pm iU)(\mathbf{n}), \quad (93)$$

where  $\mathbf{e}'_1 = \cos \psi \mathbf{e}_1 + \sin \psi \mathbf{e}_2$  and  $\mathbf{e}'_2 = -\sin \psi \mathbf{e}_1 + \cos \psi \mathbf{e}_2$ . The quantities  $Q \pm iU$  are said to be spin 2.

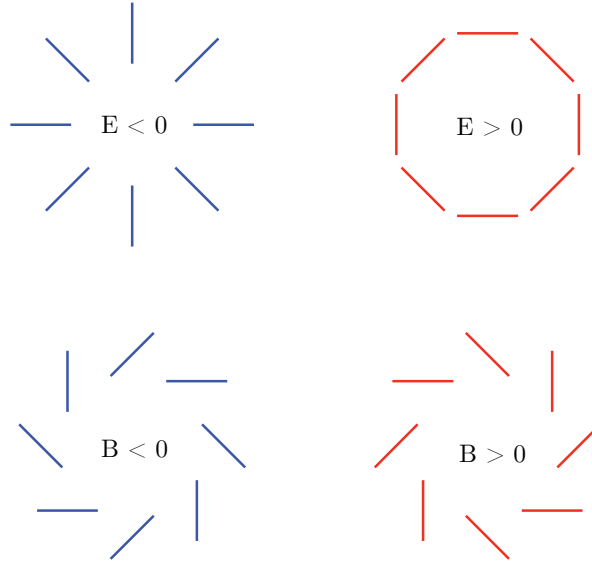
We already mentioned that the statistical properties of the radiation field are usually described in terms of the spherical harmonic decomposition of the maps. This basis, basically the Fourier basis, is very natural because the statistical properties of anisotropies are rotationally invariant. The standard spherical harmonics are not the appropriate basis for  $Q \pm iU$  because they are spin-2 variables, but generalizations (called  ${}_{\pm 2}Y_{lm}$ ) exist. We can expand

$$(Q \pm iU)(\mathbf{n}) = \sum_{\ell m} a_{\pm 2, \ell m} {}_{\pm 2}Y_{\ell m}(\mathbf{n}). \quad (94)$$

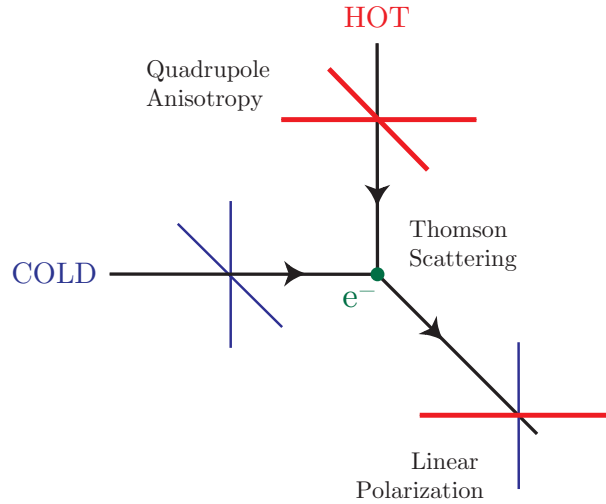
Here  $Q$  and  $U$  are defined at each direction  $\hat{\mathbf{n}}$  with respect to the spherical coordinate system ( $\mathbf{e}_\theta, \mathbf{e}_\phi$ ). To ensure that  $Q$  and  $U$  are real, the expansion coefficients must satisfy  $a_{-2, \ell m}^* = a_{2, \ell - m}$ . The equivalent relation for the temperature coefficients is  $a_{T, \ell m}^* = a_{T, \ell - m}$ . Instead of  $a_{\pm 2, \ell m}$ , it is convenient to introduce their linear combinations  $a_{E, \ell m} = -(a_{2, \ell m} + a_{-2, \ell m})/2$  and  $a_{B, \ell m} = i(a_{2, \ell m} - a_{-2, \ell m})/2$ . We define two quantities in real space,  $E(\mathbf{n}) = \sum_{\ell, m} a_{E, \ell m} Y_{\ell m}(\mathbf{n})$  and  $B(\mathbf{n}) = \sum_{\ell, m} a_{B, \ell m} Y_{\ell m}(\mathbf{n})$ . Here  $E$  and  $B$  completely specify the linear polarization field.

The temperature is a scalar quantity under a rotation of the coordinate system,  $T'(\mathbf{n}') = \mathcal{R}\mathbf{n} = T(\mathbf{n})$ , where  $\mathcal{R}$  is the rotation matrix. We denote with a prime the quantities in the transformed coordinate system. While  $Q \pm iU$  are spin 2,  $E(\mathbf{n})$  and  $B(\mathbf{n})$  are invariant under rotations. Under parity, however,  $E$  and  $B$  behave differently,  $E$  remains unchanged, while  $B$  changes sign.

To characterize the statistics of the CMB perturbations, only four power spectra are needed, those for  $T$ ,  $E$ ,  $B$  and the cross correlation between  $T$  and  $E$ . The cross correlation between  $B$  and  $E$  or  $B$  and  $T$  vanishes if there are no parity-violating interactions because  $B$  has the opposite parity to  $T$  or  $E$ . The power spectra are defined as the rotationally invariant quantities  $C_{T\ell} = \frac{1}{2\ell+1} \sum_m \langle a_{T, \ell m}^* a_{T, \ell m} \rangle$ ,  $C_{E\ell} = \frac{1}{2\ell+1} \sum_m \langle a_{E, \ell m}^* a_{E, \ell m} \rangle$ ,  $C_{B\ell} = \frac{1}{2\ell+1} \sum_m \langle a_{B, \ell m}^* a_{B, \ell m} \rangle$ , and  $C_{C\ell} = \frac{1}{2\ell+1} \sum_m \langle a_{T, \ell m}^* a_{E, \ell m} \rangle$ . The brackets  $\langle \dots \rangle$  denote ensemble averages.



**Fig. 4:** Examples of  $E$ - and  $B$ -mode patterns of polarization



**Fig. 5:** Thomson scattering of radiation where quadrupole anisotropy generates linear polarization

Polarization is generated by Thomson scattering between photons and electrons, which means that polarization cannot be generated after recombination (except for re-ionization, which we shall discuss later). But Thomson scattering is not enough. The radiation incident on the electrons must also be anisotropic. In fact, its intensity needs to have a quadrupole moment. This requirement of having both Thomson scattering and anisotropies is what makes polarization relatively small. After recombination, anisotropies grow by free streaming, but there is no scattering to generate polarization. Before recombination there were so many scatterings that they erased any anisotropy present in the photon–baryon fluid.

In the context of anisotropies induced by density perturbations, velocity gradients in the photon–baryon fluid are responsible for the quadrupole that generates polarization. Let us consider a scattering occurring at position  $\mathbf{x}_0$ : the scattered photons came from a distance of order the mean free path ( $\lambda_T$ ) away from this point. If we are considering photons traveling in direction  $\hat{\mathbf{n}}$ , they roughly come from  $\mathbf{x} = \mathbf{x}_0 + \lambda_T \hat{\mathbf{n}}$ . The photon–baryon fluid at that point was moving at velocity  $\mathbf{v}(\mathbf{x}) \approx \mathbf{v}(\mathbf{x}_0) + \lambda_T \hat{\mathbf{n}}_i \partial_i \mathbf{v}(\mathbf{x}_0)$ .

Due to the Doppler effect the temperature seen by the scatterer at  $\mathbf{x}_0$  is  $\delta T(\mathbf{x}_0, \hat{\mathbf{n}}) = \hat{\mathbf{n}} \cdot [\mathbf{v}(\mathbf{x}) - \mathbf{v}(\mathbf{x}_0)] \approx \lambda_T \hat{\mathbf{n}}_i \hat{\mathbf{n}}_j \partial_i v_j(\mathbf{x}_0)$ , which is quadratic in  $\hat{\mathbf{n}}$  (i.e., it has a quadrupole). Velocity gradients in the photon–baryon fluid lead to a quadrupole component of the intensity distribution, which, through Thomson scattering, is converted into polarization.

The polarization of the scattered radiation field, expressed in terms of the Stokes parameters  $Q$  and  $U$ , is given by  $(Q + iU) \propto \sigma_T \int d\Omega' (\mathbf{m} \cdot \hat{\mathbf{n}}')^2 T(\hat{\mathbf{n}}') \propto \lambda_p \mathbf{m}^i \mathbf{m}^j \partial_i v_j|_{\tau_{LS}}$ , where  $\sigma_T$  is the Thomson scattering cross-section and we have written the scattering matrix as  $P(\mathbf{m}, \hat{\mathbf{n}}') = -3/4 \sigma_T (\mathbf{m} \cdot \hat{\mathbf{n}}')^2$ , with  $\mathbf{m} = \hat{\mathbf{e}}_1 + i\hat{\mathbf{e}}_2$ . In the last step, we integrated over all directions of the incident photons  $\hat{\mathbf{n}}'$ . As photons decouple from the baryons, their mean free path grows very rapidly, so a more careful analysis is needed to obtain the final polarization:

$$(Q + iU)(\hat{\mathbf{n}}) \approx \epsilon \delta \tau_{LS} \mathbf{m}^i \mathbf{m}^j \partial_i v_j|_{\tau_{LS}}, \quad (95)$$

where  $\delta \tau_{LS}$  is the width of the last scattering surface and gives a measure of the distance that photons travel between their last two scatterings, and  $\epsilon$  is a numerical constant that depends on the shape of the visibility function. The appearance of  $\mathbf{m}^i \mathbf{m}^j$  in Eq. (95) ensures that  $(Q + iU)$  transforms correctly under rotations of  $(\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2)$ .

If we evaluate Eq. (95) for each Fourier mode and combine them to obtain the total power, we get the equivalent of Eq. (92),

$$\ell(\ell + 1) C_{E\ell} \approx A \epsilon^2 (1 + 3R)^2 (k_* \delta \tau_{LS})^2 \sin^2(k_* c_s \tau_{LS}), \quad (96)$$

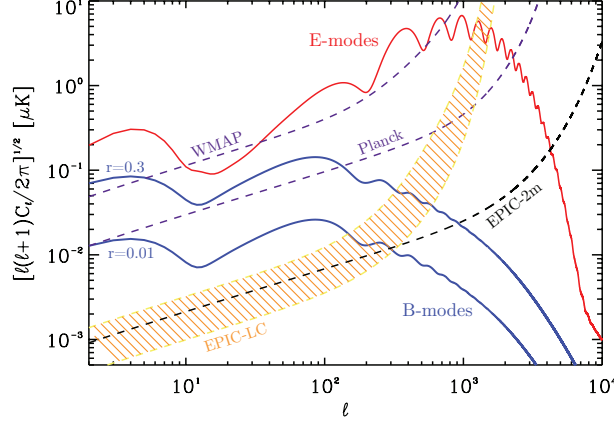
where we are assuming  $n = 1$  and that  $\ell$  is large enough that factors like  $(\ell + 2)!/(\ell - 2)! \approx \ell^4$ . The extra  $k_*$  in Eq. (96) originates in the gradient in Eq. (95). The large-angular scale polarization is greatly suppressed by the  $k \delta \tau_{LS}$  factor. Correlations over large angles can only be created by the long-wavelength perturbations, but these cannot produce a large polarization signal because of the tight coupling between photons and electrons prior to recombination. Multiple scatterings make the plasma very homogeneous; only wavelengths that are small enough to produce anisotropies over the mean free path of the photons will give rise to a significant quadrupole in the temperature distribution, and thus to polarization. Wavelengths much smaller than the mean free path decay due to photon diffusion (Silk damping) and so are unable to create a large quadrupole and polarization. As a result polarization peaks at the scale of the mean free path.

On sub-degree angular scales, temperature, polarization, and the cross-correlation power spectra show acoustic oscillations. In the polarization and cross-correlation spectra the peaks are much sharper. The polarization is produced by velocity gradients of the photon–baryon fluid at the last scatteringsurface. The temperature receives contributions from density and velocity perturbations, and the oscillations in each partially cancel one another, making the features in the temperature spectrum less sharp. The dominant contribution to the temperature comes from the oscillations in the density [Eq. (90)], which are out of phase with the velocity. This explains the difference in location between the temperature and polarization peaks. The extra gradient in the polarization signal, Eq. (95), explains why its overall amplitude peaks at a smaller angular scale.

Now, as photons travel in the metric perturbed by a GW [ $ds^2 = a^2(\tau) [-d\tau^2 + (\delta_{ij} + h_{ij}^T) dx^i dx^j]$ ], they get redshifted or blueshifted depending on their direction of propagation relative to the direction of propagation of the GW and the polarization of the GW. For example, for a GW travelling along the  $z$  axis, the frequency shift is given by

$$\frac{1}{\nu} \frac{d\nu}{d\tau} = \frac{1}{2} \hat{n}^i \hat{n}^j \dot{h}_{ij}^{T(\pm)} = \frac{1}{2} (1 - \cos^2 \theta) e^{\pm i 2\phi} \dot{h}_t \exp(i \mathbf{k} \cdot \mathbf{x}), \quad (97)$$

where  $(\theta, \phi)$  describe the direction of propagation of the photon, the  $\pm$  correspond to the different polarizations of the GW, and  $h_t$  gives the time-dependent amplitude of the GW. During the matter-dominated



**Fig. 6:** E- and B-mode power spectra for a tensor-to-scalar ratio saturating the current bounds,  $r = 0.3$  and for  $r = 0.01$ . Shown are the experimental sensitivities of WMAP, Planck and two different realizations of CMBPol (EPIC-LC and EPIC-2m)

era, for example,  $h_t = 3j_1(k\tau)/k\tau$ : time changes in the metric lead to frequency shifts (or equivalently shifts in the temperature of the black body spectrum). Notice that the angular dependence of this frequency shift is quadrupolar in nature. As a result, the temperature fluctuations induced by this effect as photons travel between successive scatterings before recombination produce a quadrupole intensity distribution, which, through Thomson scattering, lead to polarization. Both  $E$  and  $B$  power spectra are generated by GW. The current push to improve polarization measurements follows from the fact that density perturbations, to linear order in perturbation theory, cannot create any  $B$ -type polarization. As a rough rule of thumb, the amplitude of the peak in the  $B$ -mode power spectrum for GW is

$$[\ell(\ell+1)C_{Bl}/2\pi]^{1/2} = 0.024(V^{1/4}/10^{16}\text{GeV})^2\mu\text{K}$$

where

$$V^{1/4} \simeq 6.7 r^{1/4} \times 10^{16} \text{ GeV} \quad (98)$$

is the energy scale of inflation. A future experiment like CMBPol [14] can probe values of  $r$  as small as  $10^{-2}$ , corresponding to an inflation energy scale of about  $2 \times 10^{16} \text{ GeV}$ . Furthermore, using the consistency relation  $r = \epsilon$  valid in one-single field models of inflation, one deduces that

$$\frac{\Delta\phi}{m_{\text{Pl}}} \simeq \left(\frac{r}{10^{-2}}\right)^{1/2}, \quad (99)$$

meaning that a future measurement of the  $B$ -mode of CMB polarization will imply an inflaton excursions of Planckian values. Therefore, A future measurement of the  $B$ -mode polarization of the CMB will allow a determination of the value of the energy scale of inflation. This explains the utility of CMB polarization measurements as probes of the physics of inflation. A detection of primordial  $B$ -mode polarization would also demonstrate that inflation occurred at a very high energy scale, and that the inflaton traversed a super-Planckian distance in field space.

## 9 The dark puzzles

Having explored the physics of the primordial epochs of the evolution of the universe, such as inflation, and its impact on the present-day observables, we now devote the remaining space to a short discussion

of the dark puzzles of the present-day universe: the dark energy and the dark matter puzzles.

### 9.1 A present-day accelerating universe

In 1998 the accelerated expansion of the universe was pointed out by two groups from the observations of Type Ia Supernova (SN Ia) [15, 16]. Let us see how this came about.

An important concept related to observational tools in an expanding background is associated with the definition of a distance. A way of defining a distance is through the luminosity of a stellar object. The distance  $d_L$  known as the luminosity distance, plays a very important role in astronomy including in supernovae observations. It proves to be convenient to write the metric as

$$ds^2 = -dt^2 + a^2(t) [dr^2 + f_K^2(r)(d\theta^2 + \sin^2 \theta d\phi^2)] , \quad (100)$$

where

$$f_K(r) = \begin{cases} \sin r , & K = +1 , \\ r , & K = 0 , \\ \sinh r , & K = -1 . \end{cases} \quad (101)$$

In Minkowski space time the absolute luminosity  $L_s$  of the source and the energy flux  $\mathcal{F}$  at a distance  $d$  is related through  $\mathcal{F} = L_s/(4\pi d^2)$ . By generalizing this to an expanding universe, the luminosity distance,  $d_L$ , is defined as

$$d_L^2 \equiv \frac{L_s}{4\pi\mathcal{F}} . \quad (102)$$

Let us consider an object with absolute luminosity  $L_s$  located at a co-moving coordinate distance  $r$  from an observer at  $r = 0$ . The energy of light emitted from the object with time interval  $\Delta t_e$  is denoted as  $\Delta E_e$ , whereas the energy which reaches at the sphere with radius  $r$  is written as  $\Delta E_r$ . We note that  $\Delta E_e$  and  $\Delta E_r$  are proportional to the frequencies of light at  $r$  and  $r = 0$ , respectively, i.e.,  $\Delta E_e \propto \nu_e$  and  $\Delta E_r \propto \nu_r$ . The luminosities  $L_r$  and  $L_e$  are given by

$$L_r = \frac{\Delta E_e}{\Delta t_e} , \quad L_e = \frac{\Delta E_r}{\Delta t_e} . \quad (103)$$

The speed of light is given by  $c = \nu_e \lambda_e = \nu_r \lambda_r$ , where  $\lambda_e$  and  $\lambda_r$  are the wavelengths at  $r$  and  $r = 0$ . Then, we find

$$\frac{\lambda_r}{\lambda_e} = \frac{\nu_e}{\nu_r} = \frac{\Delta t_r}{\Delta t_e} = \frac{\Delta E_e}{\Delta E_r} = 1 + z , \quad (104)$$

where we have also used  $\nu_r \Delta t_r = \nu_e \Delta t_e$ . Combining Eq. (103) with Eq. (104), we obtain

$$L_e = L_r(1 + z)^2 . \quad (105)$$

The light travelling along the  $r$  direction satisfies the geodesic equation  $ds^2 = -dt^2 + a^2(t)dr^2 = 0$ . We then obtain

$$r = \int_0^r dr' = \int_{t_e}^{t_r} \frac{dt}{a(t)} \quad (106)$$

From the metric (100) we find that the area of the sphere at  $t = t_r$  is given by  $S = 4\pi(a_r f_K(r))^2$ . Hence the observed energy flux is

$$\mathcal{F} = \frac{L_r}{4\pi(a_r f_K(r))^2} . \quad (107)$$

Substituting Eqs. (106) and (107) for Eq. (102), we obtain the luminosity distance in an expanding universe:

$$d_L = a_r f_K(r)(1 + z) . \quad (108)$$



In the flat FRW background with  $f_K(r) = r$  we find

$$d_L = \left( \frac{1+z}{H_0} \right) \int_0^z dz' \frac{H_0}{H(z')}. \quad (109)$$

Then the Hubble rate  $H(z)$  can be expressed in terms of  $d_L(z)$ :

$$H(z) = \left\{ \frac{d}{dz} \left( \frac{d_L(z)}{1+z} \right) \right\}^{-1}. \quad (110)$$

If we measure the luminosity distance observationally, we can determine the expansion rate of the universe.

The energy density  $\rho$  on the right-hand side of Einstein equations includes all components present in the universe, namely, non-relativistic particles, relativistic particles, cosmological constant and so on

$$\rho = \sum_i \rho_i^{(0)} (a/a_0)^{-3(1+w_i)} = \sum_i \rho_i^{(0)} (1+z)^{3(1+w_i)}. \quad (111)$$

Here  $w_i$  and  $\rho_i^{(0)}$  correspond to the equation of state and the present energy density of each component, respectively. The Hubble parameter takes the convenient form

$$H^2 = H_0^2 \sum_i \Omega_i^{(0)} (1+z)^{3(1+w_i)}, \quad (112)$$

where  $\Omega_i^{(0)} \equiv 8\pi G \rho_i^{(0)} / (3H_0^2) = \rho_i^{(0)} / \rho_c^{(0)}$  is the density parameter for an individual component at the present epoch. Hence the luminosity distance in a flat geometry is given by

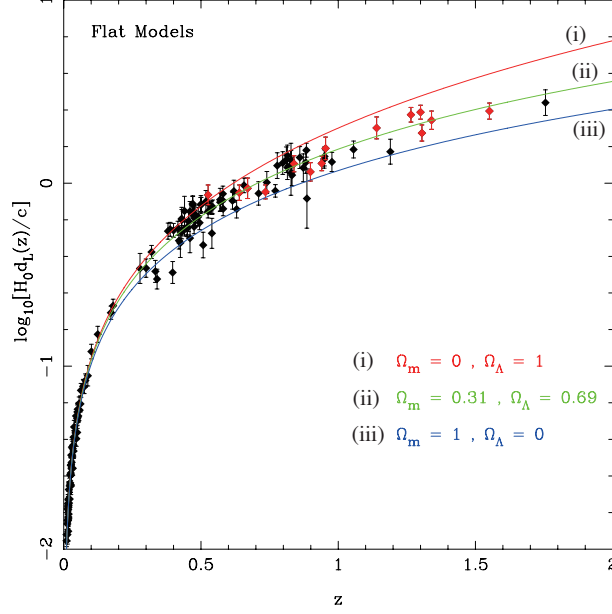
$$d_L = \frac{(1+z)}{H_0} \int_0^z \frac{dz'}{\sqrt{\sum_i \Omega_i^{(0)} (1+z')^{3(1+w_i)}}}. \quad (113)$$

The direct evidence for the current acceleration of the universe is related to the observation of luminosity distances of high redshift supernovae [15, 16]. The apparent magnitude  $m$  of the source with an absolute magnitude  $M$  is related to the luminosity distance  $d_L$  via the relation [17]

$$m - M = 5 \log_{10} \left( \frac{d_L}{\text{Mpc}} \right) + 25. \quad (114)$$

This comes from taking the logarithm of Eq. (102) by noting that  $m$  and  $M$  are related to the logarithms of  $\mathcal{F}$  and  $L_s$ , respectively. The numerical factors arise because of conventional definitions of  $m$  and  $M$  in astronomy. Type Ia supernovae (SN Ia) can be observed when white dwarf stars exceed the mass of the Chandrasekhar limit and explode. The belief is that SN Ia are formed in the same way irrespective of where they are in the universe, which means that they have a common absolute magnitude  $M$  independent of the redshift  $z$ . Thus they can be treated as an ideal standard candle. We can measure the apparent magnitude  $m$  and the redshift  $z$  observationally, which of course depends upon the objects we observe.

In order to get a feeling of the phenomenon let us consider two supernovae 1992P at low-redshift  $z = 0.026$  with  $m = 16.08$  and 1997ap at high-redshift  $z = 0.83$  with  $m = 24.32$  [15]. As we have already mentioned, the luminosity distance is approximately given by  $d_L(z) \simeq z/H_0$  for  $z \ll 1$ . Using the apparent magnitude  $m = 16.08$  of 1992P at  $z = 0.026$ , we find that the absolute magnitude is estimated by  $M = -19.09$  from Eq. (114). Here we adopted the value  $H_0^{-1} = 2998 h^{-1} \text{ Mpc}$  with  $h = 0.72$ . Then the luminosity distance of 1997ap is obtained by substituting  $m = 24.32$  and  $M = -19.09$  for Eq. (114):  $H_0 d_L \simeq 1.16$  for  $z = 0.83$ . From Eq. (113) the theoretical estimate for the luminosity distance in a two-component flat universe is  $H_0 d_L \simeq 0.95$  for  $\Omega_m^{(0)} \simeq 1$  and  $H_0 d_L \simeq 1.23$  for  $\Omega_m^{(0)} \simeq 0.3$ ,  $\Omega_{DE} \simeq 0.7$



**Fig. 7:** The luminosity distance versus redshift for a flat cosmological model. Black points are from the “Gold” data sets [18]; red points are from recent data from HST

In 2004 Riess *et al.* [18] reported the measurement of 16 high-redshift SN Ia with redshift  $z > 1.25$  with the Hubble Space Telescope (HST). By including 170 previously known SN Ia data points, they showed that the universe exhibited a transition from deceleration to acceleration at  $> 99\%$  confidence level. A best-fit value of  $\Omega_m^{(0)}$  was found to be  $\Omega_m^{(0)} = 0.29^{+0.05}_{-0.03}$  (the error bar is  $1\sigma$ ). This shows that a matter-dominated universe without a cosmological constant does not fit the data.

We should emphasize that the accelerated expansion is by cosmological standards really a late-time phenomenon, starting at a redshift  $z \sim 1$ . From Eq. (112) the deceleration parameter,  $q \equiv -a\ddot{a}/\dot{a}^2$ , is given by

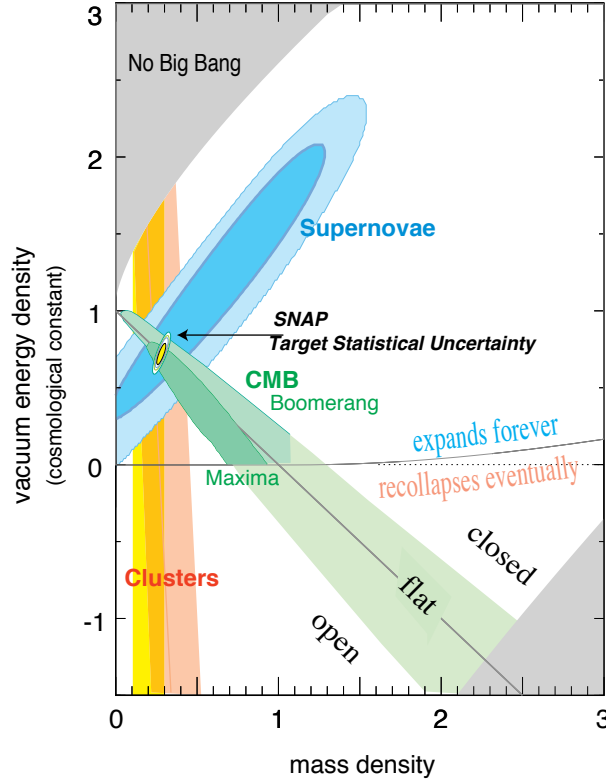
$$q(z) = \frac{3}{2} \frac{\sum_i \Omega_i^{(0)} (1 + w_i) (1 + z)^{3(1+w_i)}}{\sum_i \Omega_i^{(0)} (1 + z)^{3(1+w_i)}} - 1.$$

For the two-component flat cosmology, the universe enters an accelerating phase ( $q < 0$ ) for  $z < z_c \equiv (2\Omega_{DE}/\Omega_{DM})^{1/3} - 1$ . When  $\Omega_{DM} = 0.3$  and  $\Omega_{DE} = 0.7$ , we have  $z_c = 0.67$ . The problem of why an accelerated expansion should occur now in the long history of the universe is called the “coincidence problem”.

### 9.1.1 The origin of the acceleration

Once the idea of the accelerating universe is accepted, the next pressing question is: Why? There are various explanations available that we may mention briefly. The general trend is to accept that there is a form of Dark Energy (DE) fluid dominating the energy density of the present day. Its pressure is  $P = w\rho$  and  $w$  needs to be smaller than  $-1/3$  for this fluid to cause the acceleration. Having learned how to use scalar fields to accelerate the universe at primordial epochs, the most natural way to explain DE would be to introduce a scalar field  $\phi$  dubbed quintessence, with potential

$$\mathcal{V}(\phi) = V_0 + V(\phi), \quad (115)$$



**Fig. 8:** The dark energy (vacuum energy) and dark matter (mass density) abundances from SN, CMB, and galaxy clustering observations

Now, if  $V_0 \gg V(\phi)$  (at least at present epochs), the DE is in practice a Cosmological Constant (CC). Its value must be extremely small,  $V_0^{1/4} \simeq (H_0 m_{\text{Pl}})^{1/2} \simeq 10^{-3}$  eV. Why it is so small is a mystery that earned the name “the CC problem”. On the other hand, if  $V_0 \ll V(\phi)$ , then the dynamics of the quintessence field dominates. However, another problem arises at this stage. Having learned from inflation that the field must be slow-rolling to cause the acceleration of the universe, we have to assume that  $(m_{\text{Pl}}^2 V''/V)$  is smaller than unity. This implies that  $\phi$  is of order of the Planck scale and that its mass squared is such that  $V'' \sim H_0^2 \sim (10^{-33} \text{ eV})^2$ . The quintessence field has a Compton wavelength as large as the entire observed universe.

If the reader does not like all this fine-tuning, there are at least two other explanations for the acceleration of the universe. The first one goes under the name of modified gravity and is in fact rather intuitive. If gravity gets weaker at large distances, objects far from us may recede at a velocity larger than what they would do in the traditional Newtonian gravity case. For this to work, we have to suppose that the gravitational force has a transient at some critical (and cosmological) scale  $r_c$ , from the usual  $1/r^2$  to, say  $1/r^3$ . How to get this transition is unfortunately beyond the scope of these lectures. Another alternative goes under the name of the “anthropic principle” and is based on the following point. As we have seen, in a static universe, overdense regions will increase their density at an exponential rate. In an expanding universe, however, there is a competition between the expansion and the gravitational collapse. More rapid expansion, as induced by DE, retards the growth of structure. General relativity provides the following useful relation in linear perturbation theory between the growth factor  $g(z)$  and the expansion history of the universe

$$\ddot{g} + 2H\dot{g} = 4\pi G\rho_m = \frac{3\Omega_{\text{DM}}H_0^2}{2a^3}g. \quad (116)$$

If the universe is always matter-dominated, then  $g \sim a$ ; however, in a DE dominated universe  $g$  scales slower than the scale factor. Now, if the CC is too large, structure does not have time to develop: the initial condition is  $\delta\rho_m/\rho_m \sim 10^{-5}$  at the last scattering surface ( $z \sim 10^3$ ) and needs to become order unity by now. Now, if we impose that structures might have been able to develop by now even in the presence of a CC, one obtains a reassuring bound, the CC  $V_0^{1/4}$  must be smaller than about  $10^{-1}$  eV. In other words, the CC may not be far from the value we observe (if it is non-zero) because otherwise we would not be here to discuss about it. A great deal of observational effort of the next decades will be devoted to understand the cause of the acceleration of the universe [19]. Four observational techniques are currently receiving much attention: 1) Baryonic Acoustic Oscillations (BAO) are observed on large-scale surveys of the spatial distribution of matter. They are caused by the same oscillations that left an imprint in the CMB under the form of acoustic peaks. The BAO technique is sensitive to the DE through its effect on the angular-diameter distance vs. redshift relation and through its effect on the time evolution of the expansion rate; 2) Galaxy Cluster (CL) surveys measure the spatial density and distribution of galaxy clusters. The CL technique is sensitive to DE through its effect in the angular-diameter distance vs. redshift relation and through its effect on the time evolution of the expansion rate and the growth rate of perturbations; 3) supernovae as standard candles to determine the luminosity distance vs. redshift relation; 4) Weak Lensing (WL) surveys measure the distortion of background images due to the bending of light as it passes by galaxies or clusters of galaxies. The WL technique is sensitive to DE through its effect on the angular-diameter distance vs. redshift relation and the growth rate of perturbations. All these techniques will not only shed light on the nature of DE, but will also help us to discriminate the various possibilities to explain the present-day acceleration. For instance, the modified gravity scenario predicts a growth function which is different from the one predicted in a CC dominated universe. Future applications of the techniques briefly summarized above should be able to determine which scenario is more likely.

## 9.2 Dark matter

The evidence that 95% of the mass of galaxies and clusters is made of some unknown component of Dark Matter (DM) comes from (i) rotation curves (out to tens of kpc), (ii) gravitational lensing (out to 200 kpc), and (iii) hot gas in clusters. They lead us to believe that DM makes up about 30% of the entire energy of the universe. A nice review about DM can be found in Ref. [20].

In the 1970s, Ford and Rubin discovered that rotation curves of galaxies are flat. The velocities of objects (stars or gas) orbiting the centres of galaxies, rather than decreasing as a function of the distance from the galactic centres as had been expected, remain constant out to very large radii. Similar observations of flat rotation curves have now been found for all galaxies studied, including our Milky Way. The simplest explanation is that galaxies contain far more mass than can be explained by the bright stellar objects residing in galactic disks. This mass provides the force to speed up the orbits. To explain the data, galaxies must have enormous dark haloes made of unknown matter. Indeed, more than 95% of the mass of galaxies consists of dark matter. The baryonic matter which accounts for the gas and disk cannot alone explain the galactic rotation curve. However, adding a DM halo allows a good fit to data.

The limitations of rotation curves are that one can only look out as far as there is light or neutral hydrogen (21 cm), namely to distances of tens of kpc. Thus one can see the beginnings of DM haloes, but cannot trace where most of the DM is. The lensing experiments discussed in the next section go beyond these limitations.

Einstein's theory of General Relativity predicts that mass bends, or lenses, light. This effect can be used to gravitationally ascertain the existence of mass even when it emits no light. Lensing measurements confirm the existence of enormous quantities of DM both in galaxies and in clusters of galaxies. Observations are made of distant bright objects such as galaxies or quasars. As the result of intervening matter, the light from these distant objects is bent towards the regions of large mass. Hence there may be multiple images of the distant objects, or, if these images cannot be individually resolved, the back-

ground object may appear brighter. Some of these images may be distorted or sheared. The Sloan Digital Sky Survey used weak lensing (statistical studies of lensed galaxies) to conclude that galaxies, including the Milky Way, are even larger and more massive than previously thought, and require even more DM out to great distances. Again, the predominance of DM in galaxies is observed. The key success of the lensing of DM to date is the evidence that DM is seen out to much larger distances than could be probed by rotation curves: the DM is seen in galaxies out to 200 kpc from the centres of galaxies, in agreement with N-body simulations. On even larger Mpc scales, there is evidence for DM in filaments (the cosmic web). Another piece of gravitational evidence for DM is the hot gas in clusters. The X-ray data indicates the presence of hot gas. The existence of this gas in the cluster can only be explained by a large DM component that provides the potential well to hold on to the gas. In summary, the evidence is overwhelming for the existence of an unknown component of DM that comprises 95% of the mass in galaxies and clusters.

There is another basic reason why DM is necessary: to form structures as we observe them. Let us assume that the matter content of the universe is dominated by a pressureless and self-gravitating fluid. This approximation holds if we are dealing with the evolution of the perturbations in the DM component or in case we are dealing with structures whose size is much larger than the typical Jeans scale length of baryons. Let us also define  $\mathbf{x}$  to be the co-moving coordinate and  $\mathbf{r} = a(t)\mathbf{x}$  the proper coordinate,  $a(t)$  being the cosmic expansion factor. Furthermore, if  $\mathbf{v} = \dot{\mathbf{r}}$  is the physical velocity, then  $\mathbf{v} = \dot{a}\mathbf{x} + \mathbf{u}$ , where the first term describes the Hubble flow, while the second term,  $\mathbf{u} = a(t)\dot{\mathbf{x}}$ , gives the peculiar velocity of a fluid element which moves in an expanding background.

In this case the equations that regulate the Newtonian description of the evolution of density perturbations are the continuity equation:

$$\frac{\partial \delta}{\partial t} + \nabla \cdot [(1 + \delta)\mathbf{u}] = 0, \quad (117)$$

which gives the mass conservation, the Euler equation

$$\frac{\partial \mathbf{u}}{\partial t} + 2H(t)\mathbf{u} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\frac{\nabla \phi}{a^2}, \quad (118)$$

which gives the relation between the acceleration of the fluid element and the gravitational force, and the Poisson equation

$$\nabla^2 \phi = 4\pi G \bar{\rho} a^2 \delta \quad (119)$$

which specifies the Newtonian nature of the gravitational force. In the above equations,  $\nabla$  is the gradient computed with respect to the co-moving coordinate  $\mathbf{x}$ ,  $\phi(\mathbf{x})$  describes the fluctuations of the gravitational potential, and  $H(t) = \dot{a}/a$  is the Hubble parameter at the time  $t$ . Its time-dependence is given by  $H(t) = E(t)H_0$ , where

$$E(z) = [(1+z)^3 \Omega_m + (1+z)^2 (1 - \Omega_m - \Omega_{DE}) + (1+z)^{3(1+w)} \Omega_{DE}]^{1/2}. \quad (120)$$

In the case of small perturbations, these equations can be linearized by neglecting all the terms which are of second order in the fields  $\delta$  and  $\mathbf{u}$ . In this case, using the Euler equation to eliminate the term  $\partial \mathbf{u} / \partial t$ , and using the Poisson equation to eliminate  $\nabla^2 \phi$ , one ends up with

$$\frac{\partial^2 \delta}{\partial t^2} + 2H(t) \frac{\partial \delta}{\partial t} - 4\pi G \bar{\rho} \delta = 0. \quad (121)$$

This equation describes the Jeans instability of a pressureless fluid, with the additional ‘‘Hubble drag’’ term  $2H(t)\partial \delta / \partial t$ , which describes the counter-action of the expanding background on the perturbation growth. Its effect is to prevent the exponential growth of the gravitational instability taking place in a non-expanding background. The solution of the above equation can be cast in the form:

$$\delta(\mathbf{x}, t) = \delta_+(\mathbf{x}, t_i) D_+(t) + \delta_-(\mathbf{x}, t_i) D_-(t), \quad (122)$$

where  $D_+$  and  $D_-$  describe the growing and decaying modes of the density perturbation, respectively. In the case of an Einstein–de-Sitter (EdS) universe ( $\Omega_m = 1$ ,  $\Omega_{DE} = 0$ ), it is  $H(t) = 2/(3t)$ , so that  $D_+(t) = (t/t_i)^{2/3}$  and  $D_-(t) = (t/t_i)^{-1}$ . The fact that  $D_+(t) \propto a(t)$  for an EdS universe should not be surprising. Indeed, the dynamical time-scale for the collapse of a perturbation of uniform density  $\rho$  is  $t_{\text{dyn}} \propto (G\rho)^{-1/2}$ , while the expansion time-scale for the EdS model is  $t_{\text{exp}} \propto (G\bar{\rho})^{-1/2}$ , where  $\bar{\rho}$  is the mean cosmic density. Since for a linear (small) perturbation it is  $\rho \simeq \bar{\rho}$ , then  $t_{\text{dyn}} \sim t_{\text{exp}}$ , thus showing that the cosmic expansion and the perturbation evolution take place at the same pace. This argument also leads to understanding the behaviour for a  $\Omega_m < 1$  model. In this case, the expansion time scale becomes shorter than the above one at the redshift at which the universe recognizes that  $\Omega_m < 1$ . This happens at  $1 + z \simeq \Omega_m^{-1/3}$  or at  $1 + z \simeq \Omega_m^{-1}$  in the presence or absence of a cosmological constant term, respectively. Therefore, after this redshift, cosmic expansion takes place at a quicker pace than gravitational instability, with the result that the perturbation growth is frozen.

The exact expression for the growing model of perturbations is given by

$$D_+(z) = \frac{5}{2} \Omega_m E(z) \int_z^\infty \frac{1+z'}{E(z')^3} dz'. \quad (123)$$

The EdS has the faster evolution, while the slowing down of the perturbation growth is more apparent for the open low-density model, the presence of a cosmological constant providing an intermediate degree of evolution. The key point is, however, that a pressureless fluid such as DM is needed for the perturbations to grow to give rise to collapsed objects. Baryon perturbations, being coupled to photons till the last-scattering epoch, feel a non-vanishing pressure and therefore they may not grow. After the last-scattering stage, the baryons fall into the gravitational potential generated by DM and the baryonic perturbations may promptly catch up with those of DM.

### 9.2.1 Dark matter candidates

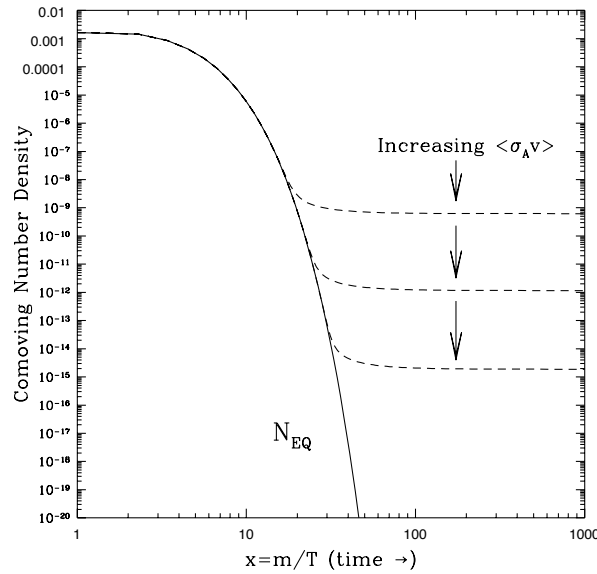
There is a plethora of dark matter candidates. MACHOs, or Massive Compact Halo Objects, are made of ordinary matter in the form of faint stars or stellar remnants; they could also be primordial black holes or mirror matter. However, there are not enough of them to completely resolve the question. Of the non-baryonic candidates, the most popular are the WIMPS (Weakly Interacting Massive Particles) and the axions, as these particles have been proposed for other reasons in particle physics. Ordinary massive neutrinos are too light to be cosmologically significant, though sterile neutrinos remain a possibility. Other candidates include primordial black holes, non-thermal WIMPzillas, and Kaluza–Klein particles which arise in higher dimensional theories.

About axions, the good news is that cosmologists do not need to “invent” new particles. Two candidates already exist in particle physics for other reasons: axions and WIMPs. Axions with masses in the range  $10^{-(3-6)}$  eV arise in the Peccei–Quinn solution to the strong-CP problem in the theory of strong interactions.

WIMPs are also natural dark matter candidates from particle physics. These particles, if present in thermal abundances in the early universe, annihilate with one another so that a predictable number of them remain today. The relic density of these particles comes out to be the right value:

$$\Omega_{\text{DM}} h^2 = (3 \times 10^{-26} \text{ cm}^3/\text{s}) / \langle \sigma v \rangle_A \quad (124)$$

where the annihilation cross-section  $\langle \sigma v \rangle_A$  of weak interaction strength automatically gives the right answer. The reason why the final abundance is inversely proportional to the annihilation cross-section is rather clear: the larger the annihilation cross-section, the more WIMPs annihilate and the fewer of them are left behind. Furthermore, annihilation is not eternal: owing to the expansion of the universe, annihilation stops when its rate becomes smaller than the expansion rate of the universe. When this happens, the abundance is said to freeze-out.



**Fig. 9:** The abundance of WIMPs of a given mass  $m$  as a function of temperature and for various annihilation cross-sections

This coincidence is known as ‘the WIMP miracle’ and is the reason why WIMPs are taken so seriously as DM candidates. The best WIMP candidate is motivated by Supersymmetry (SUSY): the lightest neutralino in the Minimal Supersymmetric Standard Model. Supersymmetry in particle theory is designed to keep particle masses at the right value. As a consequence, each particle we know has a partner: the photino is the partner of the photon, the squark is the quark’s partner, and the selectron is the partner of the electron. The lightest supersymmetric partner is a good dark matter candidate.

There are several ways to search for dark WIMPs. SUSY particles may be discovered at the LHC as missing energy in an event. In that case one knows that the particles live long enough to escape the detector, but it will still be unclear whether they are long-lived enough to be the dark matter. Thus complementary astrophysical experiments are needed. In direct detection experiments, the WIMP scatters off a nucleus in the detector, and a number of experimental signatures of the interaction can be detected. In indirect detection experiments, neutrinos that arise as annihilation products of captured WIMPs exit from the Sun and can be detected on Earth. Another way to detect WIMPs is to look for anomalous cosmic rays from the Galactic Halo: WIMPs in the Halo can annihilate with one another to give rise to antiprotons, positrons, or neutrinos. In addition, neutrinos, gamma rays, and radio waves may be detected as WIMP annihilation products from the Galactic Centre. For lack of time these issues were not discussed extensively in the lectures. The interested reader may find more about these issues in Ref. [20].

## 10 Conclusions

The period when we say that cosmology is entering a golden age has already passed: cosmology *is* in the middle of its golden age. Present observational data pose various puzzles whose solutions might either be around the corner or decades far in the future. It will require some young and creative researcher sitting in this room to solve them. This is why the cosmological puzzles are dark, but the future is brighter.

## Acknowledgements

It is a great pleasure to thank all the organizers, N. Ellis, E. Lillostol, D. Metral, and especially M. Losada and E. Nardi, for having created such a stimulating atmosphere. All students are also acknowledged for their never-ending enthusiasm.

## References

- [1] A. D. Linde, *Particle Physics and Inflationary Cosmology* (Harwood, Chur, Switzerland, 1990) and references therein.
- [2] E. W. Kolb and M. S. Turner, *The Early Universe* (Addison-Wesley, Redwood City, CA, 1990).
- [3] A. R. Liddle and D. H. Lyth, Phys. Rep. **231**, 1 (1993) and references therein.
- [4] A. R. Liddle and D. H. Lyth, *Cosmological Inflation and Large-Scale Structure* (Cambridge University Press, 2000) and references therein.
- [5] D. H. Lyth and A. Riotto, Phys. Rep. **314**, 1 (1999) and references therein.
- [6] A. Riotto, arXiv:hep-ph/0210162 and references therein.
- [7] For a review, see N. Bartolo, S. Matarrese, and A. Riotto, arXiv:astro-ph/0703496 and references therein.
- [8] E. J. Copeland, M. Sami, and S. Tsujikawa, Int. J. Mod. Phys. D **15**, 1753 (2006) and references therein.
- [9] E. Komatsu *et al.* [WMAP Collaboration], Astrophys. J. Suppl. **180**, 330 (2009) [arXiv:0803.0547 [astro-ph]].
- [10] For a review, see, for instance, N. A. Bahcall, J. P. Ostriker, S. Perlmutter, and P. J. Steinhardt, Science **284**, 1481 (1999).
- [11] R. K. Sachs and A. M. Wolfe, Astrophys. J. **147**, 73 (1967).
- [12] S. Dodelson, *Modern Cosmology* (Academic Press, New York, 2003).
- [13] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger, Phys. Rep. **215**, 203 (1992) and references therein.
- [14] S. Dodelson *et al.*, CMBPol Science White Paper submitted to the US Astro2010 Decadal Survey, arXiv:0902.3796v1.
- [15] S. Perlmutter *et al.*, Astrophys. J. **517**, 565 (1999).
- [16] A. G. Riess *et al.*, Astron. J. **116**, 1009 (1998); Astron. J. **117**, 707 (1999).
- [17] T. Padmanabhan, Phys. Rep. **380**, 235 (2003); T. Padmanabhan, Current Science, **88**, 1057 (2005) [arXiv:astro-ph/0510492].
- [18] A. G. Riess *et al.* [Supernova Search Team Collaboration], Astrophys. J. **607**, 665 (2004).
- [19] A. J. Albrecht *et al.*, arXiv:astro-ph/0609591.
- [20] G. Bertone, D. Hooper, and J. Silk, Phys. Rep. **405**, 279 (2005).